

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2020年10月15日(15.10.2020)



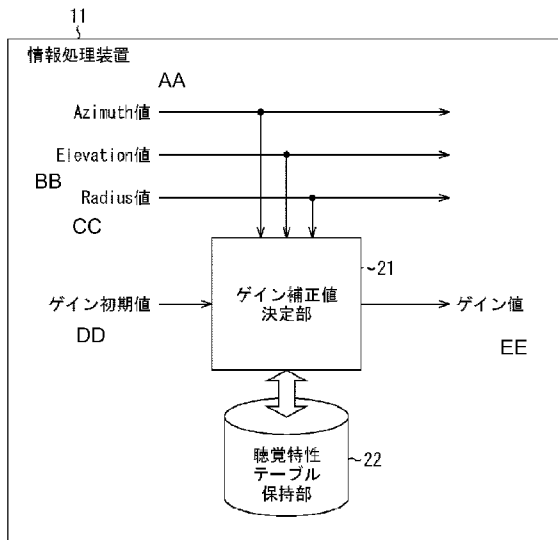
(10) 国際公開番号
WO 2020/209103 A1

- (51) 国際特許分類:
H04R 3/00 (2006.01) *H04S 7/00* (2006.01)
H04S 5/00 (2006.01)
- (21) 国際出願番号: PCT/JP2020/014120
- (22) 国際出願日: 2020年3月27日(27.03.2020)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2019-075369 2019年4月11日(11.04.2019) JP
- (71) 出願人: ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒1080075 東京都港区港南1丁目7番1号 Tokyo (JP).
- (72) 発明者: 辻 実 (TSUJI Minoru); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP). 知念 徹 (CHINEN Toru); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP). 山本 優樹 (YAMAMOTO Yuki); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP). 中井 彬人 (NAKAI Akihito); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP).
- (74) 代理人: 西川 孝, 外 (NISHIKAWA Takashi et al.); 〒1700013 東京都豊島区東池袋3丁目9番10号 池袋F Nビル4階 Tokyo (JP).
- (81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ,

(54) Title: INFORMATION PROCESSING DEVICE AND METHOD, REPRODUCTION DEVICE AND METHOD, AND PROGRAM

(54) 発明の名称: 情報処理装置および方法、再生装置および方法、並びにプログラム

FIG. 4



- 11 Information processing device
- 21 Gain correction value determination unit
- 22 Aural characteristic table storage unit
- AA Azimuth value
- BB Elevation value
- CC Radius value
- DD Initial gain value
- EE Gain value

(57) Abstract: This technology relates to an information processing device and a method, a reproduction device and a method, and a program which are configured to enable gain correction to be carried out more easily. This information processing device includes a gain correction value determination unit that determines a correction value of a gain value for gain correcting an audio signal of an audio object according to the direction of the audio object as viewed by an audience. The present technology can be applied to a gain determination device and a reproduction device.



WO 2020/209103 A1

BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

- (84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類 :

- 一 国際調査報告 (条約第21条(3))

(57) 要約 : 本技術は、より簡単にゲイン補正を行うことができるようにする情報処理装置および方法、再生装置および方法、並びにプログラムに関する。情報処理装置は、聴取者から見たオーディオオブジェクトの方向に応じて、オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正值を決定するゲイン補正值決定部を備える。本技術はゲイン決定装置や再生装置に適用することができる。

明 細 書

発明の名称：

情報処理装置および方法、再生装置および方法、並びにプログラム

技術分野

[0001] 本技術は、情報処理装置および方法、再生装置および方法、並びにプログラムに関し、特に、より簡単にゲイン補正を行うことができるようにした情報処理装置および方法、再生装置および方法、並びにプログラムに関する。

背景技術

[0002] 従来、MPEG (Moving Picture Experts Group) -H 3D Audio規格が知られている（例えば、非特許文献1および非特許文献2参照）。

[0003] MPEG-H 3D Audio規格等で扱われる3D Audioでは、3次元的な音の方向や距離、拡がりなどを再現することができ、従来のステレオ再生に比べ、より臨場感のあるオーディオ再生が可能となる。

先行技術文献

非特許文献

[0004] 非特許文献1：ISO/IEC 23008-3, MPEG-H 3D Audio

非特許文献2：ISO/IEC 23008-3:2015/AMENDMENT3, MPEG-H 3D Audio Phase 2

発明の概要

発明が解決しようとする課題

[0005] しかしながら3D Audioでは、コンテンツ（3D Audioコンテンツ）の制作の時間的なコストが高くなってしまふ。

[0006] 例えば3D Audioでは、ステレオと比較してオブジェクトの位置情報、すなわち音源の位置情報の次元数が高い（3D Audioは3次元でステレオは2次元）。そのため、3D Audioでは、特にオブジェクトの位置を示す水平角度や垂直角度、距離、オブジェクトについてのゲインなどといったオブジェクトごとのメタデータを構成するパラメタを決定する作業において、時間的なコストが高くなってしまふ。

[0007] また、3D Audioコンテンツはステレオコンテンツに比べて、コンテンツと制作者の両面で圧倒的に数が少ない。それゆえ、品質の高い3D Audioコンテンツが少ないのが現状である。

[0008] 一方で、聴覚特性として、音の大きさの感じ方は、その音の到来方向によって異なる。すなわち、同じオブジェクトの音であっても、オブジェクトが聴取者に対して前方にある場合と側方にある場合、上方にある場合と下方にある場合で、それぞれ聴感上の音の大きさが異なるため、このような聴覚特性を踏まえたゲイン補正が必要である。

[0009] 以上のことから、より簡単にゲイン補正を行い、これにより、短時間で十分な品質の3D Audioコンテンツを制作できるようにすることが望まれている。

[0010] 本技術は、このような状況に鑑みてなされたものであり、より簡単にゲイン補正を行うことができるようにするものである。

課題を解決するための手段

[0011] 本技術の第1の側面の情報処理装置は、聴取者から見たオーディオオブジェクトの方向に応じて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正值を決定するゲイン補正值決定部を備える。

[0012] 本技術の第1の側面の情報処理方法またはプログラムは、聴取者から見たオーディオオブジェクトの方向に応じて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正值を決定するステップを含む。

[0013] 本技術の第1の側面においては、聴取者から見たオーディオオブジェクトの方向に応じて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正值が決定される。

[0014] 本技術の第2の側面の再生装置は、オーディオオブジェクトの位置を示す位置情報に基づいて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正值であって、聴取者から見た前記オーディオ

オオブジェクトの方向に応じた補正値を決定し、前記補正値により補正された前記ゲイン値に基づいて前記オーディオ信号の前記ゲイン補正を行うゲイン補正部と、前記ゲイン補正により得られた前記オーディオ信号に基づいてレンダリング処理を行い、前記オーディオオブジェクトの音を再生するための複数のチャンネルの再生信号を生成するレンダラ処理部とを備える。

[0015] 本技術の第2の側面の再生方法またはプログラムは、オーディオオブジェクトの位置を示す位置情報に基づいて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正値であって、聴取者から見た前記オーディオオブジェクトの方向に応じた補正値を決定し、前記補正値により補正された前記ゲイン値に基づいて前記オーディオ信号の前記ゲイン補正を行い、前記ゲイン補正により得られた前記オーディオ信号に基づいてレンダリング処理を行い、前記オーディオオブジェクトの音を再生するための複数のチャンネルの再生信号を生成するステップを含む。

[0016] 本技術の第2の側面においては、オーディオオブジェクトの位置を示す位置情報に基づいて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正値であって、聴取者から見た前記オーディオオブジェクトの方向に応じた補正値が決定され、前記補正値により補正された前記ゲイン値に基づいて前記オーディオ信号の前記ゲイン補正が行われ、前記ゲイン補正により得られた前記オーディオ信号に基づいてレンダリング処理が行われ、前記オーディオオブジェクトの音を再生するための複数のチャンネルの再生信号が生成される。

図面の簡単な説明

[0017] [図1]音の到来方向に対する聴覚特性について説明する図である。

[図2]音の到来方向に対する聴覚特性について説明する図である。

[図3]音の到来方向に対する聴覚特性について説明する図である。

[図4]情報処理装置の構成例を示す図である。

[図5]聴覚特性テーブルの例を示す図である。

[図6]聴覚特性テーブルの例を示す図である。

- [図7]ゲイン値決定処理を説明するフローチャートである。
- [図8]コンテンツ制作ツールの表示画面例を示す図である。
- [図9]コンテンツ制作ツールの表示画面例を示す図である。
- [図10]コンテンツ制作ツールの表示画面例を示す図である。
- [図11]コンテンツ制作ツールの表示画面例を示す図である。
- [図12]情報処理装置の構成例を示す図である。
- [図13]テーブル生成処理を説明するフローチャートである。
- [図14]音声処理装置の構成例を示す図である。
- [図15]再生信号生成処理を説明するフローチャートである。
- [図16]聴覚特性テーブルの例を示す図である。
- [図17]ゲイン聴覚特性情報のシンタックス例を示す図である。
- [図18]音声処理装置の構成例を示す図である。
- [図19]コンピュータの構成例を示す図である。

発明を実施するための形態

[0018] 以下、図面を参照して、本技術を適用した実施の形態について説明する。

[0019] 〈第1の実施の形態〉

〈本技術について〉

本技術は聴取者から見たオブジェクトの方向に応じてゲイン補正値を決定することで、より簡単にゲイン補正を行うことができるようにし、これにより、より簡単に、すなわち短時間で十分に高い品質の3D Audioコンテンツを制作できるようにするものである。

[0020] 特に、本技術は以下の特徴（F1）乃至特徴（F5）を有している。

[0021] 特徴（F1）：オブジェクトのゲイン補正値を、音像の定位位置に対する3次元聴覚特性に応じて決定する

特徴（F2）：聴覚特性がテーブル等により与えられる場合、データの無い定位位置に対するゲイン補正値は、隣接位置のゲイン補正値に基づく補間処理等により算出する

特徴（F3）：自動ミキシングにおいて、別途決定した位置情報からゲイ

ン情報を決定する

特徴（F4）：オブジェクト位置に対するゲイン補正値を設定および調整するユーザインターフェースを提供する

特徴（F5）：聴取位置に対するオブジェクトの位置の変更に伴い、3次元聴覚特性に応じたゲイン補正値を適用する

[0022] まず、人間の3次元聴覚特性に基づくゲインパラメタの決定について説明する。

[0023] 図1は、あるピンクノイズが聴取者の真正面で再生されたときの聴感上の音の大きさを基準として、同じピンクノイズを異なる方向から再生した際に、聴感上の音の大きさが同じに感じるようにピンクノイズのゲイン補正を行ったときのゲイン補正量を示している。換言すれば、図1は人が有する水平方向に対する聴覚特性を示している。

[0024] なお、図1において縦軸はゲイン補正量を示し、横軸は聴取者から見た音源位置を示す水平方向の角度であるAzimuth値（水平角度）を示している。

[0025] 例えば、聴取者から見た真正面の方向を示すAzimuth値は0度であり、聴取者から見た真横の方向、つまり側方を示すAzimuth値は±90度であり、聴取者の後方、つまり真後ろの方向を示すAzimuth値は180度である。特に、聴取者から見て左方向がAzimuth値の正の方向となっている。

[0026] また、図1ではピンクノイズの再生時の垂直方向の位置は、聴取者と同じ高さとなっている。すなわち、聴取者から見た音源の垂直方向（仰角方向）の位置を示す垂直角度をElevation値とすると、図1はElevation値が0度における場合の例となっている。なお、聴取者から見て上方向がElevation値の正の方向となっている。

[0027] この例では、複数人の聴取者を対象として行われた実験の結果から得られた各Azimuth値に対するゲイン補正量の平均値を示しており、特に各Azimuth値において点線で表される範囲は95%の信頼区間を示している。

[0028] 例えば側方（Azimuth値＝±90度、Elevation値＝0度）でピンクノイズを再生するときには、ゲインを少し下げることによって、聴取者は正面方向でピンクノ

イズを再生したときと同じ大きさで音が聞こえるように感じる事が分かる。

[0029] また、例えば後方（Azimuth値=180度，Elevation値=0度）でピンクノイズを再生するときには、ゲインを少し上げることで、聴取者は正面方向でピンクノイズを再生したときと同じ大きさで音が聞こえるように感じる事が分かる。

[0030] すなわち、あるオブジェクト音源に対して、そのオブジェクト音源の定位位置が聴取者の側方にあるときにはオブジェクト音源の音のゲインを少し下げ、オブジェクト音源の定位位置が聴取者の後方にあるときにはオブジェクト音源の音のゲインを少し上げると、聴取者に同じ大きさで音が聞こえているように感じさせることができる。

[0031] また、例えば図2や図3に示すように、同じAzimuth値でもElevation値が変化すると、聴取者の聞こえ方も変化することが分かる。

[0032] なお、図2および図3において縦軸はゲイン補正量を示し、横軸は聴取者から見た音源位置を示すAzimuth値（水平角度）を示している。また、図2および図3では、各Azimuth値において点線で表される範囲は95%の信頼区間を示している。

[0033] 図2は、Elevation値が30度である場合における各Azimuth値でのゲイン補正量を示している。

[0034] 図2から、音源が聴取者よりも高い位置にある場合、音源が聴取者の正面や後方、斜め後方にあるときには音が小さく聞こえ、音源が聴取者の斜め前方にあるときには音が少し大きく聞こえることが分かる。

[0035] 同様に、図3はElevation値が-30度である場合における各Azimuth値でのゲイン補正量を示している。

[0036] 図3から、音源が聴取者よりも低い位置にある場合、音源が聴取者の正面や斜め前方にあるときには音が大きく聞こえ、音源が聴取者の後方や斜め後方にあるときには音が小さく聞こえることが分かる。

[0037] 以上のような音の到来方向に対する聴覚特性から、オブジェクト音源の位

置を示す位置情報と、聴取者の聴覚特性とに基づいて、オブジェクト音源に対するゲイン補正量を決定すれば、より簡単に適切なゲイン補正を行うことができることが分かる。

[0038] 〈情報処理装置の構成例〉

図4は、本技術を適用した情報処理装置の一実施の形態の構成例を示す図である。

[0039] 図4に示す情報処理装置11は、3D Audioコンテンツを構成するオーディオオブジェクト（以下、単にオブジェクトと称する）の音を再生するためのオーディオ信号のゲイン補正のためのゲイン値を決定するゲイン決定装置として機能する。

[0040] このような情報処理装置11は、例えば3D Audioコンテンツを構成するオーディオ信号のミキシングを行う編集装置などに設けられている。

[0041] 情報処理装置11は、ゲイン補正值決定部21および聴覚特性テーブル保持部22を有している。

[0042] ゲイン補正值決定部21には、3D Audioコンテンツを構成するオブジェクトのメタデータとして、位置情報およびゲイン初期値が供給される。

[0043] ここで、オブジェクトの位置情報は、3次元空間内における基準位置から見たオブジェクトの位置を示す情報であり、ここでは位置情報はAzimuth値、Elevation値、およびRadius値からなる。なお、この例では聴取者の位置が基準位置となっている。

[0044] Azimuth値およびElevation値は、基準位置にいる聴取者（ユーザ）から見たオブジェクトの水平方向および垂直方向の各位置を示す角度であり、これらのAzimuth値およびElevation値は図1乃至図3における場合と同様である。

[0045] また、Radius値は3次元空間における基準位置にいる聴取者からオブジェクトまでの距離（半径）である。

[0046] このようなAzimuth値、Elevation値、およびRadius値からなる位置情報は、オブジェクトの音の音像の定位位置を示しているということができる。

- [0047] また、ゲイン補正值決定部 2 1 に供給されるメタデータに含まれるゲイン初期値は、オブジェクトのオーディオ信号のゲイン補正のためのゲイン値、つまりゲイン情報の初期値であり、このゲイン初期値は、例えば 3D Audio コンテンツの制作者等により定められる。なお、ここでは説明を簡単にするため、ゲイン初期値は 1.0 であるものとする。
- [0048] ゲイン補正值決定部 2 1 は、供給されたメタデータとしての位置情報と、聴覚特性テーブル保持部 2 2 に保持されている聴覚特性テーブルとに基づいて、オブジェクトのゲイン初期値を補正するゲイン補正量を示すゲイン補正值を決定する。
- [0049] また、ゲイン補正值決定部 2 1 は、決定したゲイン補正值に基づいて、供給されたゲイン初期値を補正し、その結果得られたゲイン値を、オブジェクトのオーディオ信号をゲイン補正するための最終的なゲイン補正量を示す情報とする。
- [0050] 換言すれば、ゲイン補正值決定部 2 1 は位置情報により示される、聴取者から見たオブジェクトの方向（音の到来方向）に応じてゲイン補正值を決定することで、オーディオ信号のゲイン値を決定する。このようにして決定されたゲイン値と、供給された位置情報とがオブジェクトの最終的なメタデータとして後段に出力される。
- [0051] 聴覚特性テーブル保持部 2 2 は、聴覚特性テーブルを保持しており、必要に応じて聴覚特性テーブルにより示されるゲイン補正值をゲイン補正值決定部 2 1 に供給する。
- [0052] ここで、聴覚特性テーブルは、音源であるオブジェクトから聴取者への音の到来方向、つまり聴取者から見た音源の方向と、その方向に応じたゲイン補正值とが対応付けられているテーブルである。
- [0053] すなわち、より詳細には、聴覚特性テーブルは音源と聴取者との相対的な位置関係と、その位置関係に応じたゲイン補正值とが対応付けられているテーブルである。
- [0054] 聴覚特性テーブルにより示されるゲイン補正值は、例えば図 1 乃至図 3 に

示したような音の到来方向に対する人の聴覚特性に応じて定められたものであり、特に音の到来方向によらず聴感上の音の大きさが一定になるようなゲイン補正量となっている。

[0055] すなわち、聴覚特性テーブルにより示されるゲイン補正值によりゲイン初期値を補正して得られたゲイン値を用いてオブジェクトのオーディオ信号をゲイン補正すれば、オブジェクトの位置によらず、同じオブジェクトの音は同じ大きさで聞こえるようになる。

[0056] ここで、図5に聴覚特性テーブルの例を示す。

[0057] 図5に示す例ではAzimuth値、Elevation値、およびRadius値により定まるオブジェクトの位置、つまりオブジェクトの方向に対してゲイン補正值が対応付けられている。

[0058] 特に、この例では全てのElevation値およびRadius値が0および1.0となっており、オブジェクトの垂直方向の位置は聴取者と同じ高さであり、かつ聴取者からオブジェクトまでの距離は常に一定であると想定されている。

[0059] 図5の例では、例えばAzimuth値が180度である場合など、音源であるオブジェクトが聴取者の後方にある場合には、Azimuth値が0度や30度である場合など、オブジェクトが聴取者の前方にある場合よりもゲイン補正值が大きくなっている。

[0060] これに対して、例えばAzimuth値が90度である場合など、音源であるオブジェクトが聴取者の側方にある場合には、オブジェクトが聴取者の前方にある場合よりもゲイン補正值が小さくなっている。

[0061] さらに、聴覚特性テーブル保持部22が図5に示す聴覚特性テーブルを保持している場合における、ゲイン補正值決定部21によるゲイン初期値の補正の具体的な例について説明する。

[0062] 例えばオブジェクトの位置を示すAzimuth値、Elevation値、およびRadius値が90度、0度、および1.0mであるとすると、図5からオブジェクトの位置に対応するゲイン補正值は-0.52dBとなる。

[0063] したがって、ゲイン補正值決定部21は、聴覚特性テーブルから読み出し

たゲイン補正值「-0.52dB」と、ゲイン初期値「1.0」とに基づいて次式（1）の計算を行い、ゲイン値「0.94」を得る。

[0064] [数1]

$$1.0 \times 10^{-0.52/20} \cong 0.94 \quad \dots (1)$$

[0065] 同様に、例えばオブジェクトの位置を示すAzimuth値、Elevation値、およびRadius値が-150度、0度、および1.0mであるとする、図5からオブジェクトの位置に対応するゲイン補正值は0.51dBとなる。

[0066] したがって、ゲイン補正值決定部21は、聴覚特性テーブルから読み出したゲイン補正值「0.51dB」と、ゲイン初期値「1.0」とに基づいて次式（2）の計算を行い、ゲイン値「1.06」を得る。

[0067] [数2]

$$1.0 \times 10^{0.51/20} \cong 1.06 \quad \dots (2)$$

[0068] なお、図5では水平方向のみが考慮された2次元の聴覚特性に基づいて決定されたゲイン補正值を利用する例について説明した。つまり、2次元の聴覚特性に基づいて生成された聴覚特性テーブル（以下、2次元聴覚特性テーブルとも称する）を利用する例について説明した。

[0069] しかし、水平方向だけでなく垂直方向の特性も考慮された3次元の聴覚特性に基づいて決定されたゲイン補正值を利用してゲイン初期値を補正するようによい。

[0070] そのような場合、例えば図6に示す聴覚特性テーブルを利用することができる。

[0071] 図6に示す例では、Azimuth値、Elevation値、およびRadius値により定まるオブジェクトの位置、つまりオブジェクトの方向に対してゲイン補正值が対応付けられている。

[0072] 特に、この例では全てのAzimuth値およびElevation値の組み合わせにおいて、Radius値は1.0とされている。

[0073] 以下では、図6に示すように音の到来方向に対する3次元の聴覚特性に基

づいて生成された聴覚特性テーブルを、特に3次元聴覚特性テーブルとも称することとする。

[0074] ここで、聴覚特性テーブル保持部22が図6に示す聴覚特性テーブルを保持している場合における、ゲイン補正值決定部21によるゲイン初期値の補正の具体的な例について説明する。

[0075] 例えばオブジェクトの位置を示すAzimuth値、Elevation値、およびRadius値が60度、30度、および1.0mであるとする、図6からオブジェクトの位置に対応するゲイン補正值は-0.07dBとなる。

[0076] したがって、ゲイン補正值決定部21は、聴覚特性テーブルから読み出したゲイン補正值「-0.07dB」と、ゲイン初期値「1.0」とに基づいて次式(3)の計算を行い、ゲイン値「0.99」を得る。

[0077] [数3]

$$1.0 \times 10^{-0.07/20} \cong 0.99 \quad \dots (3)$$

[0078] なお、以上において説明したゲイン値算出の具体例では、オブジェクトの位置(方向)に対して定まる聴覚特性に基づくゲイン補正值が予め用意されていた。すなわち、オブジェクトの位置情報に対応するゲイン補正值が聴覚特性テーブルに格納されている例について説明した。

[0079] しかしながら、オブジェクトの位置は、聴覚特性テーブルにおいて、対応するゲイン補正值が格納されている位置にあるとは限らない。

[0080] 具体的には、例えば聴覚特性テーブル保持部22に図6に示した聴覚特性テーブルが保持されており、位置情報としてのAzimuth値、Elevation値、およびRadius値が-120度、15度、および1.0mであるとする。

[0081] この場合、図6の聴覚特性テーブルには、Azimuth値「-120」、Elevation値「15」、およびRadius値「1.0」に対応するゲイン補正值は格納されていない。

[0082] そこで、聴覚特性テーブルに、位置情報により示される位置に対応するゲイン補正值がない場合には、その位置情報により示される位置に隣接する、対応するゲイン補正值が存在する複数の位置のデータ(ゲイン補正值)を用

いて、ゲイン補正值決定部 21 が補間処理等により所望の位置のゲイン補正值を算出するようにしてもよい。

[0083] 換言すれば、聴取者から見たオブジェクトの方向（位置）に対応するゲイン補正值が聴覚特性テーブルに格納されていない場合には、そのゲイン補正值を、聴取者から見たオブジェクトの他の方向に対応するゲイン補正值に基づく補間処理等により求めてもよい。

[0084] 例えば、ゲイン補正值の補間方法の 1 つとして VBAP (Vector Base Amplitude Panning) がある。

[0085] VBAP は、オブジェクトごとに、オブジェクトのメタデータから再生環境の複数のスピーカのゲイン値を求めるためのものである。

[0086] ここで、再生環境の複数のスピーカを、複数のゲイン補正值に置き換えることで、所望の位置でのゲイン補正值を算出することができる。

[0087] 具体的には、3次元空間内においてゲイン補正值が用意されている複数の位置でメッシュが区切られる。すなわち、例えば3次元空間内の3つの各位置のゲイン補正值が用意されているとすると、それらの3つの位置を頂点とする1つの3角形の領域が1つのメッシュとされる。

[0088] このようにして3次元空間が複数のメッシュに区切られると、ゲイン補正值を得ようとする所望の位置を注目位置として、その注目位置を内包するメッシュが特定される。

[0089] また、特定されたメッシュを構成する3つの頂点位置を示す位置ベクトルの乗加算により注目位置を示す位置ベクトルを表したときの3つの各頂点位置を示す位置ベクトルに乘算される係数が求められる。

[0090] そして、このようにして求められた3つの係数のそれぞれが、注目位置を内包するメッシュの3つの各頂点位置のゲイン補正值のそれぞれに乘算され、係数が乘算されたゲイン補正值の和が注目位置のゲイン補正值として算出される。

[0091] 具体的には、注目位置を内包するメッシュの3つの各頂点位置を示す位置ベクトルが P_1 乃至 P_3 であり、それらの各頂点位置のゲイン補正值が G_1 乃至 G_3 で

あるとする。

[0092] このとき、注目位置を示す位置ベクトルが $g_1P_1+g_2P_2+g_3P_3$ で表されたとする。
この場合、注目位置のゲイン補正值は、 $g_1G_1+g_2G_2+g_3G_3$ となる。

[0093] なお、ゲイン補正值の補間方法はVBAPによる補間に限らず、他のどのような方法であってもよい。

[0094] 例えば聴覚特性テーブルにおいてゲイン補正值が存在する位置のうち、注目位置の近傍にあるN個（例えばN=5など）の位置のゲイン補正值の平均値を注目位置のゲイン補正值として用いてもよい。

[0095] また、例えば聴覚特性テーブルにおいてゲイン補正值が存在する位置のうち、注目位置から最も近い位置のゲイン補正值を注目位置のゲイン補正值として用いてもよい。

[0096] さらに、ここではゲイン補正值がデシベル値で求められる例について説明したが、ゲイン補正值はリニア値で求められるようにしてもよい。そのような場合、例えばVBAPによる補間によってリニア値でゲイン補正值を求めるときでも、上述したデシベル値における場合と同様の計算により任意の位置のゲイン補正值を得ることができる。

[0097] その他、オブジェクトの種類や優先度、音圧、音高などに基づいて、そのオブジェクトのメタデータとしての位置情報、すなわちAzimuth値、Elevation値、およびRadius値を決定する場合にも本技術は適用可能である。

[0098] この場合、例えばオブジェクトの種類や優先度等に基づいて決定された位置情報と、予め用意された3次元聴覚特性テーブルとに基づいてゲイン補正值が決定される。

[0099] 〈ゲイン値決定処理の説明〉

続いて、情報処理装置11の動作について説明する。すなわち、以下、図7のフローチャートを参照して、情報処理装置11により行われるゲイン値決定処理について説明する。

[0100] ステップS11においてゲイン補正值決定部21は、外部からメタデータを取得する。

- [0101] すなわち、ゲイン補正值決定部 2 1 はAzimuth値、Elevation値、およびRadius値からなる位置情報とゲイン初期値とをメタデータとして取得する。
- [0102] ステップS 1 2においてゲイン補正值決定部 2 1は、ステップS 1 1で取得した位置情報と、聴覚特性テーブル保持部 2 2に保持されている聴覚特性テーブルとに基づいてゲイン補正值を決定する。
- [0103] すなわち、ゲイン補正值決定部 2 1は、取得した位置情報を構成するAzimuth値、Elevation値、およびRadius値に対応付けられているゲイン補正值を聴覚特性テーブルから読み出して、読み出したゲイン補正值を決定したゲイン補正值とする。
- [0104] ステップS 1 3においてゲイン補正值決定部 2 1は、ステップS 1 1で取得したゲイン初期値と、ステップS 1 2で決定したゲイン補正值とに基づいてゲイン値を決定する。
- [0105] すなわち、ゲイン補正值決定部 2 1は、ゲイン初期値およびゲイン補正值に基づいて式(1)と同様の計算を行ってゲイン初期値をゲイン補正值により補正することで、ゲイン値を得る。
- [0106] このようにしてゲイン値が決定されると、ゲイン補正值決定部 2 1は、決定されたゲイン値を後段に出力し、ゲイン値決定処理は終了する。出力されたゲイン値は、後段においてオーディオ信号のゲイン補正(ゲイン調整)に用いられる。
- [0107] 以上のようにして情報処理装置 1 1は、聴覚特性テーブルを用いてゲイン補正值を決定し、そのゲイン補正值によりゲイン初期値を補正することでゲイン値を決定する。
- [0108] このようにすることで、より簡単にゲイン補正を行うことができる。これにより、例えばより簡単に、すなわち短時間で十分に高い品質の3D Audioコンテンツを制作できるようになる。
- [0109] <第2の実施の形態>
<ユーザインターフェースについて>
また、本技術によれば、以上において説明したゲイン補正值を設定したり

調整したりするためのユーザインターフェースを提供することが可能である。

[0110] 例えばユーザの入力により、または自動でオブジェクトの位置等を決定する3Dオーディオのコンテンツ制作ツールに本技術を適用することができる。

[0111] 具体的には、3Dオーディオのコンテンツ制作ツールにおいて、例えば図8に示すユーザインターフェース（表示画面）により、聴取者から見たオブジェクトの方向に対する聴覚特性に基づくゲイン補正值（ゲイン値）の設定や調整を行うようにすることができる。

[0112] 図8に示す例では、3Dオーディオのコンテンツ制作ツールの表示画面には、予めプリセットされた互いに異なる複数の聴覚特性のなかから所望の聴覚特性を選択するためのプルダウンボックスBX11が設けられている。

[0113] この例では、例えば男性の聴覚特性、女性の聴覚特性、ユーザ個人の聴覚特性など、複数の2次元の聴覚特性が予め用意されており、ユーザはプルダウンボックスBX11を操作することで、所望の聴覚特性を選択することができる。

[0114] ユーザにより聴覚特性が選択されると、図中、プルダウンボックスBX11の下側に設けられたゲイン補正值表示領域R11に、ユーザにより選択された聴覚特性に応じた各Azimuth値でのゲイン補正值が表示される。

[0115] 特に、ゲイン補正值表示領域R11において縦軸はゲイン補正值を示しており、横軸はAzimuth値を示している。

[0116] また、曲線L11はAzimuth値が負の値、つまり聴取者から見て右方向についての各Azimuth値におけるゲイン補正值を示しており、曲線L12は聴取者から見て左方向についての各Azimuth値におけるゲイン補正值を示している。

[0117] このようなゲイン補正值表示領域R11を見れば、ユーザは直感的かつ瞬時に各Azimuth値におけるゲイン補正值を把握することができる。

[0118] さらに、ゲイン補正值表示領域R11の図中、下側には、ゲイン補正值表示領域R11に表示されたゲイン補正值を調整するためのスライダ等が表示されたスライダ表示領域R12が設けられている。

- [0119] スライド表示領域R12では、ユーザがゲイン補正値を調整可能な各Azimuth値について、そのAzimuth値を示す数字と、ゲイン補正値を示す目盛り、およびゲイン補正値を調整するためのスライドが表示されている。
- [0120] 例えばスライドSD11は、Azimuth値が30度におけるゲイン補正値を調整するためのものであり、ユーザはスライドSD11を上下に移動させることで、調整後のゲイン補正値として所望の値を指定することができる。
- [0121] スライドSD11によりゲイン補正値が調整されると、その調整に応じてゲイン補正値表示領域R11の表示が更新される。すなわち、ここではスライドSD11に対する操作に応じて曲線L12が変化する。
- [0122] このように図8に示す例では、聴取者から見た右側の各方向のゲイン補正値と、聴取者から見た左側の各方向のゲイン補正値とを独立して調整することが可能である。
- [0123] 特に、この例では予め用意された複数の聴覚特性のなかから任意のものを選択することで、所望の聴覚特性に応じたゲイン補正値、つまり聴覚特性テーブルを指定することができる。そして、スライドを操作することで、選択した聴覚特性に応じたゲイン補正値をさらに調整することができる。
- [0124] 例えば、予め用意された聴覚特性は平均的なものであるため、ユーザはスライドを操作することで、ユーザ個人の聴覚特性に応じたものとなるようにゲイン補正値を調整することができる。また、スライドを操作してゲイン補正値を調整すれば、後方のオブジェクトについては大きめのゲイン補正をして強調させるなどといったユーザの意図に沿った調整を行うこともできるようになる。
- [0125] このようにして各Azimuth値におけるゲイン補正値が設定および調整され、例えば図示せぬ保存ボタン等が操作されると、ゲイン補正値表示領域R11に表示されたゲイン補正値と、各Azimuth値とが対応付けられた2次元聴覚特性テーブルが生成される。
- [0126] なお、図8では、聴取者から見た右側と左側の各方向でゲイン補正値が異なる例、つまりゲイン補正値が左右非対称である例について説明した。しか

し、ゲイン補正值は左右対称となるようにしてもよい。

[0127] そのような場合、例えば図9に示すようにゲイン補正值の設定や調整が行われる。なお、図9において図8における場合と対応する部分には同一の符号を付してあり、その説明は適宜省略する。

[0128] 図9は、3Dオーディオのコンテンツ制作ツールの表示画面を示しており、この例では表示画面にはプルダウンボックスBX11、ゲイン補正值表示領域R21、およびスライダ表示領域R22が表示されている。

[0129] ゲイン補正值表示領域R21には、図8のゲイン補正值表示領域R11と同様に各Azimuth値でのゲイン補正值が表示されるが、ここでは左右の各方向のゲイン補正值が共通であるため、ゲイン補正值を示す曲線が1つだけ表示されている。

[0130] 例えば、左右の各方向のゲイン補正值の平均値を、左右で共通化されたゲイン補正值とするなどとすることができる。この場合、例えば図8の例におけるAzimuth値が90度および-90度のそれぞれのゲイン補正值の平均値が、図9の例におけるAzimuth値が±90度の共通のゲイン補正值とされる。

[0131] また、スライダ表示領域R22には、ゲイン補正值表示領域R21に表示されたゲイン補正值を調整するためのスライダ等が表示されている。

[0132] 例えば、この例ではユーザはスライダSD21を上下に移動させることで、Azimuth値が±30度の共通のゲイン補正值を調整することができる。

[0133] さらに、例えば図10に示すようにElevation値ごとに各Azimuth値でのゲイン補正值を調整できるようにしてもよい。なお、図10において図8における場合と対応する部分には同一の符号を付してあり、その説明は適宜省略する。

[0134] 図10は3Dオーディオのコンテンツ制作ツールの表示画面を示しており、この例では表示画面にはプルダウンボックスBX11、ゲイン補正值表示領域R31乃至ゲイン補正值表示領域R33、およびスライダ表示領域R34乃至スライダ表示領域R36が表示されている。

[0135] 図10に示す例では、図9に示した例と同様にゲイン補正值が左右対称と

なっている。

[0136] ゲイン補正值表示領域R31には、Elevation値が30度であるときの各Azimuth値でのゲイン補正值が表示されており、ユーザはスライダ表示領域R34に表示されたスライダ等进行操作することで、それらのゲイン補正值を調整することができる。

[0137] 同様に、ゲイン補正值表示領域R32には、Elevation値が0度であるときの各Azimuth値でのゲイン補正值が表示されており、ユーザはスライダ表示領域R35に表示されたスライダ等进行操作することで、それらのゲイン補正值を調整することができる。

[0138] また、ゲイン補正值表示領域R33には、Elevation値が-30度であるときの各Azimuth値でのゲイン補正值が表示されており、ユーザはスライダ表示領域R36に表示されたスライダ等进行操作することで、それらのゲイン補正值を調整することができる。

[0139] このようにして各Azimuth値におけるゲイン補正值が設定および調整され、例えば図示せぬ保存ボタン等が操作されると、ゲイン補正值とElevation値およびAzimuth値とが対応付けられた3次元聴覚特性テーブルが生成される。

[0140] さらに、3Dオーディオのコンテンツ制作ツールの表示画面の他の例として、図11に示すようにレーダーチャート型のゲイン補正值表示領域が設けられるようにしてもよい。なお、図11において図10における場合と対応する部分には同一の符号を付してあり、その説明は適宜省略する。

[0141] 図11の例では表示画面にはプルダウンボックスBX11、ゲイン補正值表示領域R41乃至ゲイン補正值表示領域R43、およびスライダ表示領域R34乃至スライダ表示領域R36が表示されている。この例では、図10に示した例と同様にゲイン補正值が左右対称となっている。

[0142] ゲイン補正值表示領域R41には、Elevation値が30度であるときの各Azimuth値でのゲイン補正值が表示されており、ユーザはスライダ表示領域R34に表示されたスライダ等进行操作することで、それらのゲイン補正值を調整することができる。

- [0143] 特にゲイン補正值表示領域R41では、レーダーチャートの各項目がAzimuth値となっているので、ユーザは各方向（Azimuth値）とそれらの方向のゲイン補正值だけでなく、各方向間のゲイン補正值の相対的な差も瞬時に把握することができる。
- [0144] ゲイン補正值表示領域R41と同様に、ゲイン補正值表示領域R42には、Elevation値が0度であるときの各Azimuth値でのゲイン補正值が表示されている。また、ゲイン補正值表示領域R43には、Elevation値が-30度であるときの各Azimuth値でのゲイン補正值が表示されている。
- [0145] 〈情報処理装置の構成例〉
- 次に、図8等を参照して説明した3Dオーディオのコンテンツ制作ツールにより聴覚特性テーブルを生成する情報処理装置について説明する。
- [0146] そのような情報処理装置は、例えば図12に示すように構成される。
- [0147] 図12に示される情報処理装置51は、コンテンツ制作ツールを実現し、そのコンテンツ制作ツールの表示画面を表示装置52に表示させる。
- [0148] 情報処理装置51は、入力部61、聴覚特性テーブル生成部62、聴覚特性テーブル保持部63、および表示制御部64を有している。
- [0149] 入力部61は、例えばマウスやキーボード、スイッチ、ボタン、タッチパネル等からなり、ユーザの操作に応じた入力信号を聴覚特性テーブル生成部62に供給する。
- [0150] 聴覚特性テーブル生成部62は、入力部61から供給された入力信号、および聴覚特性テーブル保持部63に保持されているプリセットされた聴覚特性の聴覚特性テーブルに基づいて新たな聴覚特性テーブルを生成し、聴覚特性テーブル保持部63に供給する。
- [0151] また、聴覚特性テーブル生成部62は、聴覚特性テーブル生成時に適宜、表示制御部64に対して表示装置52における表示画面の表示の更新等を指示する。
- [0152] 聴覚特性テーブル保持部63は、予めプリセットされた聴覚特性の聴覚特性テーブルを保持しており、その聴覚特性テーブルを適宜、聴覚特性テーブ

ル生成部62に供給するとともに、聴覚特性テーブル生成部62から供給された聴覚特性テーブルを保持する。

[0153] 表示制御部64は、聴覚特性テーブル生成部62の指示に従って表示装置52による表示画面の表示を制御する。

[0154] なお、図12に示した入力部61、聴覚特性テーブル生成部62、および表示制御部64が図4に示した情報処理装置11に設けられるようにしてもよい。

[0155] 〈テーブル生成処理の説明〉

続いて、情報処理装置51の動作について説明する。

[0156] すなわち、以下、図13のフローチャートを参照して、情報処理装置51により行われるテーブル生成処理について説明する。

[0157] ステップS41において表示制御部64は、聴覚特性テーブル生成部62の指示に応じて、表示装置52にコンテンツ制作ツールの表示画面を表示させる。

[0158] 具体的には、例えば表示制御部64は、表示装置52に図8や図9、図10、図11などに示した表示画面を表示させる。

[0159] このとき、例えばユーザが入力部61を操作してプリセットされた聴覚特性を選択した場合には、聴覚特性テーブル生成部62は、入力部61から供給される入力信号に応じて、ユーザにより選択された聴覚特性に対応する聴覚特性テーブルを聴覚特性テーブル保持部63から読み出す。

[0160] そして聴覚特性テーブル生成部62は、読み出した聴覚特性テーブルにより示される各Azimuth値のゲイン補正值が表示装置52に表示されるように、ゲイン補正值表示領域の表示を表示制御部64に指示する。表示制御部64は、聴覚特性テーブル生成部62の指示に応じて、表示装置52における表示画面にゲイン補正值表示領域を表示させる。

[0161] 表示装置52にコンテンツ制作ツールの表示画面が表示されると、ユーザは適宜、入力部61を操作して、スライダ表示領域に表示されたスライダ等を操作することにより、ゲイン補正值の変更（調整）を指示する。

- [0162] すると、ステップS 4 2において聴覚特性テーブル生成部6 2は、入力部6 1から供給された入力信号に応じて聴覚特性テーブルを生成する。
- [0163] すなわち、聴覚特性テーブル生成部6 2は、聴覚特性テーブル保持部6 3から読み出した聴覚特性テーブルを、入力部6 1から供給された入力信号に応じて変更することで、新たな聴覚特性テーブルを生成する。すなわち、プリセットされた聴覚特性テーブルが、スライダ表示領域に表示されたスライダ等の操作に応じて変更（更新）される。
- [0164] このようにしてスライダ等の操作に応じて各Azimuth値のゲイン補正值が調整（変更）され、新たな聴覚特性テーブルが生成されると、聴覚特性テーブル生成部6 2は、その新たな聴覚特性テーブルに応じてゲイン補正值表示領域の表示の更新を表示制御部6 4に指示する。
- [0165] ステップS 4 3において表示制御部6 4は、聴覚特性テーブル生成部6 2の指示に従って表示装置5 2を制御し、新たに生成された聴覚特性テーブルに応じた表示を行う。
- [0166] 具体的には、表示制御部6 4は、表示装置5 2における表示画面上のゲイン補正值表示領域の表示を、新たに生成された聴覚特性テーブルに応じて更新させる。
- [0167] ステップS 4 4において聴覚特性テーブル生成部6 2は、入力部6 1から供給される入力信号に基づいて、処理を終了するか否かを判定する。
- [0168] 例えば聴覚特性テーブル生成部6 2は、ユーザが入力部6 1を操作して、表示装置5 2に表示されている保存ボタン等を操作することで、入力信号として聴覚特性テーブルの保存を指示する旨の信号が供給された場合、処理を終了すると判定する。
- [0169] ステップS 4 4において、まだ処理を終了しないと判定された場合、処理はステップS 4 2に戻り、上述した処理が繰り返し行われる。
- [0170] これに対して、ステップS 4 4において処理を終了すると判定された場合、処理はステップS 4 5に進む。
- [0171] ステップS 4 5において聴覚特性テーブル生成部6 2は、最後に行ったス

トップS 4 2 で得られた聴覚特性テーブルを、新たに生成された聴覚特性テーブルとして聴覚特性テーブル保持部 6 3 に供給し、保持させる。

[0172] 聴覚特性テーブル保持部 6 3 に聴覚特性テーブルが保持されると、テーブル生成処理は終了する。

[0173] 以上のようにして情報処理装置 5 1 は、表示装置 5 2 にコンテンツ制作ツールの表示画面を表示させ、ユーザの操作に応じてゲイン補正値を調整することで、新たな聴覚特性テーブルを生成する。

[0174] このようにすることで、ユーザは簡単かつ直感的に所望の聴覚特性に応じた聴覚特性テーブルを得ることができる。したがって、ユーザは、より簡単に、すなわち短時間で十分に高い品質の3D Audioコンテンツを制作することができる。

[0175] 〈第3の実施の形態〉

〈音声処理装置の構成例〉

また、例えば自由視点のコンテンツでは、3次元空間における聴取者の位置を自由に移動させることができるため、聴取者の移動に伴って3次元空間におけるオブジェクトと聴取者の相対的な位置関係も変化する。

[0176] このように聴取者の位置を自由に移動させることができる場合に、聴取者の位置の変更に応じて音源位置を補正し、その結果得られた補正位置情報に基づいてレンダリング処理を行う技術が提案されている（例えば、国際公開第2015/107926号参照）。

[0177] 本技術は、このような自由視点のコンテンツを再生する再生装置にも適用可能である。そのような場合、補正位置情報だけでなく、上述の3次元の聴覚特性も用いられてゲイン補正が行われる。

[0178] 図14は、本技術を適用した、自由視点のコンテンツを再生する再生装置として機能する音声処理装置の一実施の形態の構成例を示す図である。なお、図14において図4における場合と対応する部分には同一の符号を付しており、その説明は適宜省略する。

[0179] 図14に示す音声処理装置91は、入力部121、位置情報補正部122

、ゲイン／周波数特性補正部 1 2 3、聴覚特性テーブル保持部 2 2、空間音響特性付加部 1 2 4、レンダラ処理部 1 2 5、および畳み込み処理部 1 2 6 を有している。

[0180] 音声処理装置 9 1 には、再生対象となるコンテンツのオーディオ情報として、オブジェクトごとに、オブジェクトのオーディオ信号と、オーディオ信号のメタデータとが供給される。なお、図 1 4 では、情報処理装置 9 1 に 2 つのオブジェクトのオーディオ信号およびメタデータが供給される例について説明するが、これに限らずオブジェクトの数はいくつであってもよい。

[0181] ここで、音声処理装置 9 1 に供給されるメタデータは、オブジェクトの位置情報およびゲイン初期値である。

[0182] また、位置情報は上述したAzimuth値、Elevation値、およびRadius値からなり、3次元空間内における基準位置から見たオブジェクトの位置、すなわちオブジェクトの音の定位位置を示す情報である。なお、以下、3次元空間における基準位置を、特に標準聴取位置とも称することとする。

[0183] 入力部 1 2 1 はマウスやボタン、タッチパネルなどからなり、ユーザにより操作されると、その操作に応じた信号を出力する。例えば入力部 1 2 1 は、ユーザによる想定聴取位置の入力を受け付け、ユーザにより入力された想定聴取位置を示す想定聴取位置情報を位置情報補正部 1 2 2 および空間音響特性付加部 1 2 4 に供給する。

[0184] ここで、想定聴取位置は、再現したい仮想の音場における、コンテンツを構成する音の聴取位置である。したがって、想定聴取位置は、予め定められた標準聴取位置を変更（補正）したときの変更後の位置を示しているということができる。

[0185] 位置情報補正部 1 2 2 は、入力部 1 2 1 から供給された想定聴取位置情報と、外部から供給された聴取者の向きを示す方向情報とに基づいて、外部から供給されたオブジェクトのメタデータとしての位置情報を補正する。

[0186] 位置情報補正部 1 2 2 は、位置情報の補正により得られた補正位置情報をゲイン／周波数特性補正部 1 2 3 およびレンダラ処理部 1 2 5 に供給する。

- [0187] なお、方向情報は、例えばユーザ（聴取者）の頭部に設けられたジャイロセンサ等から得ることができる。また、補正位置情報は、想定聴取位置において、方向情報により示される方向を向いている聴取者から見たオブジェクトの位置、つまりオブジェクトの音の定位位置を示す情報である。
- [0188] ゲイン／周波数特性補正部 1 2 3 は、位置情報補正部 1 2 2 から供給された補正位置情報と、聴覚特性テーブル保持部 2 2 に保持されている聴覚特性テーブルと、外部から供給されたメタデータとに基づいて、外部から供給されたオブジェクトのオーディオ信号のゲイン補正および周波数特性補正を行う。
- [0189] ゲイン／周波数特性補正部 1 2 3 は、ゲイン補正および周波数特性補正により得られたオーディオ信号を空間音響特性付加部 1 2 4 に供給する。
- [0190] 空間音響特性付加部 1 2 4 は、入力部 1 2 1 から供給された想定聴取位置情報と、外部から供給されたオブジェクトの位置情報とに基づいて、ゲイン／周波数特性補正部 1 2 3 から供給されたオーディオ信号に空間音響特性を付加し、レンダラ処理部 1 2 5 に供給する。
- [0191] レンダラ処理部 1 2 5 は、位置情報補正部 1 2 2 から供給された補正位置情報に基づいて、空間音響特性付加部 1 2 4 から供給されたオーディオ信号に対するレンダリング処理、すなわちマッピング処理を行い、2 以上である M 個のチャンネルの再生信号を生成する。
- [0192] すなわち、各オブジェクトのオーディオ信号から、 M チャンネルの再生信号が生成される。レンダラ処理部 1 2 5 は、生成された M チャンネルの再生信号を畳み込み処理部 1 2 6 に供給する。
- [0193] このようにして得られた M チャンネルの再生信号は、仮想的な M 個のスピーカ（ M チャンネルのスピーカ）で再生することで、再現したい仮想の音場の想定聴取位置において聴取される、各オブジェクトから出力された音を再現するオーディオ信号である。
- [0194] 畳み込み処理部 1 2 6 は、レンダラ処理部 1 2 5 から供給された M チャンネルの再生信号に対する畳み込み処理を行い、2チャンネルの再生信号を生成

して出力する。

[0195] すなわち、この例ではコンテンツの再生側の機器はヘッドホンとされており、畳み込み処理部126では、ヘッドホンに設けられた2つのスピーカ（ドライバ）で再生される再生信号が生成され、出力される。

[0196] 〈再生信号生成処理の説明〉

続いて、音声処理装置91の動作について説明する。

[0197] すなわち、以下、図15のフローチャートを参照して、音声処理装置91により行われる再生信号生成処理について説明する。

[0198] ステップS71において入力部121は、想定聴取位置の入力を受け付ける。

[0199] 入力部121は、ユーザが入力部121を操作して想定聴取位置を入力すると、その想定聴取位置を示す想定聴取位置情報を位置情報補正部122および空間音響特性付加部124に供給する。

[0200] ステップS72において位置情報補正部122は、入力部121から供給された想定聴取位置情報と、外部から供給されたオブジェクトの位置情報および方向情報とに基づいて補正位置情報を算出する。

[0201] 位置情報補正部122は、各オブジェクトについて得られた補正位置情報を、ゲイン／周波数特性補正部123およびレンダラ処理部125に供給する。

[0202] ステップS73において、ゲイン／周波数特性補正部123は、位置情報補正部122から供給された補正位置情報と、外部から供給されたメタデータと、聴覚特性テーブル保持部22に保持されている聴覚特性テーブルとに基づいて、外部から供給されたオブジェクトのオーディオ信号のゲイン補正および周波数特性補正を行う。

[0203] 具体的には、例えばゲイン／周波数特性補正部123は、聴覚特性テーブルから、補正位置情報を構成するAzimuth値、Elevation値、およびRadius値に対応付けられているゲイン補正值を読み出す。

[0204] また、ゲイン／周波数特性補正部123は、メタデータとして供給された

位置情報のRadius値と、補正位置情報のRadius値との比をゲイン補正值に乘算することでゲイン補正值を補正し、その結果得られたゲイン補正值によりゲイン初期値を補正してゲイン値を得る。

[0205] これにより、想定聴取位置から見たオブジェクトの方向に応じたゲイン補正と、想定聴取位置からオブジェクトまでの距離に応じたゲイン補正とがゲイン値によるゲイン補正によって実現されることになる。

[0206] さらにゲイン／周波数特性補正部123は、メタデータとして供給された位置情報のRadius値と、補正位置情報のRadius値とに基づいてフィルタ係数を選択する。

[0207] このようにして選択されたフィルタ係数は、所望の周波数特性補正を実現するためのフィルタ処理に用いられる。より具体的には、例えばフィルタ係数は、想定聴取位置からオブジェクトまでの距離に応じて、再現したい仮想の音場の壁や天井によって、オブジェクトからの音の高域成分が減衰する特性を再現するためのものである。

[0208] ゲイン／周波数特性補正部123は、以上のようにして得られたフィルタ係数とゲイン値に基づいて、オブジェクトのオーディオ信号に対するゲイン補正およびフィルタ処理を行うことで、ゲイン補正と周波数特性補正を実現する。

[0209] ゲイン／周波数特性補正部123は、ゲイン補正および周波数特性補正により得られた各オブジェクトのオーディオ信号を空間音響特性付加部124に供給する。

[0210] ステップS74において空間音響特性付加部124は、入力部121から供給された想定聴取位置情報と、外部から供給されたオブジェクトの位置情報とに基づいて、ゲイン／周波数特性補正部123から供給されたオーディオ信号に空間音響特性を付加し、レンダラ処理部125に供給する。

[0211] 例えば空間音響特性付加部124は、オブジェクトの位置情報と想定聴取位置情報とから定まる遅延量およびゲイン量に基づいて、オーディオ信号に対してマルチタップディレイ処理やコムフィルタ処理、オールパスフィルタ

処理を施すことで、空間音響特性の付加を行う。これにより、例えば空間音響特性として初期反射や残響特性などがオーディオ信号に付加される。

[0212] ステップS 7 5においてレンダラ処理部 1 2 5は、位置情報補正部 1 2 2から供給された補正位置情報に基づいて、空間音響特性付加部 1 2 4から供給されたオーディオ信号に対するマッピング処理を行うことで、Mチャンネルの再生信号を生成し、畳み込み処理部 1 2 6に供給する。

[0213] 例えばステップS 7 5の処理では、VBAPにより再生信号が生成されるが、その他、どのような手法でMチャンネルの再生信号が生成されるようにしてもよい。

[0214] ステップS 7 6において畳み込み処理部 1 2 6は、レンダラ処理部 1 2 5から供給されたMチャンネルの再生信号に対する畳み込み処理を行うことで、2チャンネルの再生信号を生成し、出力する。例えば畳み込み処理としてBRI R (Binaural Room Impulse Response) 処理が行われる。

[0215] 2チャンネルの再生信号が生成されて出力されると、再生信号生成処理は終了する。

[0216] 以上のようにして音声処理装置 9 1は、想定聴取位置情報に基づいて補正位置情報を算出するとともに、得られた補正位置情報や想定聴取位置情報に基づいて、各オブジェクトのオーディオ信号のゲイン補正や周波数特性補正を行ったり、空間音響特性を付加したりする。

[0217] これにより、より簡単に適切なゲイン補正や周波数特性補正を行うことができる。また、各オブジェクトから出力された音の任意の想定聴取位置での聞こえ方をリアルに再現することができる。したがって、ユーザはコンテンツの再生時に自身の嗜好に合わせて、自由に聴取位置を指定することができるようになり、より自由度の高いオーディオ再生を実現することができる。

[0218] なお、ステップS 7 3では、補正位置情報に基づいて、想定聴取位置からオブジェクトまでの距離に応じたゲイン補正と周波数特性補正が行われることに加え、聴覚特性テーブルが用いられて、3次元の聴覚特性に基づいたゲイン補正も行われる。

- [0219] このとき、ステップS 7 3 で用いられる聴覚特性テーブルは、例えば図 1 6 に示すものなどとされる。
- [0220] 図 1 6 に示す聴覚特性テーブルは、図 6 に示した聴覚特性テーブルにおけるゲイン補正値の符号を反転させることで得られたものとなっている。
- [0221] このような聴覚特性テーブルを用いてゲイン初期値を補正すれば、同じオブジェクト（音源）であっても、そのオブジェクトからの音の到来方向によって聴感上の音の大きさが変化するという現象をゲイン補正により再現することができる。これにより、よりリアリティの高い音場再現を実現することができる。
- [0222] 一方で、再生条件によっては図 1 6 に示した聴覚特性テーブルよりも、図 6 に示した聴覚特性テーブルを用いる方がより適切なゲイン補正を実現できることもある。
- [0223] すなわち、例えばコンテンツの再生にヘッドホンが用いられるのではなく、3次元空間に配置された実スピーカを用いたスピーカ再生が行われる場合について考える。
- [0224] この場合、音声処理装置 9 1 では、レンダラ処理部 1 2 5 により得られたMチャンネルの再生信号が、それらのM個の各チャンネルに対応するスピーカに供給されてコンテンツの音が再生されることになる。
- [0225] このような実スピーカを用いたコンテンツ再生では、実際に音源、つまりオブジェクトの音が想定聴取位置から見たオブジェクトの位置で再生される。
- [0226] そのため、音の到来方向によって聴感上の音の大きさが変化するという現象を再現するようなゲイン補正は不要であり、むしろ音量バランスを変えないように、聴感上の音の大きさを変化させたくないこともある。
- [0227] そのようなときには、ステップS 7 3 において図 6 に示した聴覚特性テーブルを用いてゲイン補正値を決定し、そのゲイン補正値を用いてゲイン初期値を補正すればよい。そうすれば、オブジェクトのある方向によらず、聴感上の音の大きさが一定になるようなゲイン補正が行われる。

[0228] 〈第3の実施の形態の変形例1〉

〈ゲイン聴覚特性情報の符号伝送について〉

ところで、オーディオ信号やメタデータなどが符号化されて符号化ビットストリームにより伝送されることがある。

[0229] そのような場合、例えばゲイン／周波数特性補正部123において、聴覚特性テーブルを用いたゲイン補正を行うか否かのフラグ情報等が含まれたゲイン聴覚特性情報を符号化ビットストリームにより伝送することもできる。

[0230] このとき、ゲイン聴覚特性情報にはフラグ情報だけでなく、聴覚特性テーブルや、複数の聴覚特性テーブルのうちのゲイン補正に用いる聴覚特性テーブルを示すインデックス情報なども含まれるようにすることができる。

[0231] このようなゲイン聴覚特性情報のシンタックスは、例えば図17に示すようなものとすることができる。

[0232] 図17の例では、文字「numGainAuditoryPropertyTables」は、符号化ビットストリームにより伝送する聴覚特性テーブルの数、つまりゲイン聴覚特性情報に含まれている聴覚特性テーブルの数を示している。

[0233] また、文字「numElements[i]」は、ゲイン聴覚特性情報に含まれているi番目の聴覚特性テーブルを構成する要素の数を示している。

[0234] ここでいう要素とは、互いに対応付けられたAzimuth値、Elevation値、Radius値、およびゲイン補正值である。

[0235] さらに文字「azimuth[i][n]」、「elevation[i][n]」、および「radius[i][n]」は、i番目の聴覚特性テーブルのn番目の要素を構成するAzimuth値、Elevation値、およびRadius値を示している。

[0236] 換言すれば、azimuth[i][n]、elevation[i][n]、およびradius[i][n]は、音源であるオブジェクトの音の到来方向、つまりオブジェクトの位置を示す水平角度、垂直角度、および距離（半径）を示している。

[0237] また、文字「gainCompensValue[i][n]」は、i番目の聴覚特性テーブルのn番目の要素を構成するゲイン補正值、すなわちazimuth[i][n]、elevation[i][n]、およびradius[i][n]により示される位置（方向）に対するゲイン補正值

を示している。

[0238] さらに、文字「hasGainCompensObjects」は、聴覚特性テーブルを用いたゲイン補正を行うオブジェクトがあるか否かを示すフラグ情報である。

[0239] また、文字「num_objects」は、コンテンツを構成するオブジェクトの数（オブジェクト数）を示しており、このオブジェクト数num_objectsは、ゲイン聴覚特性情報とは別にコンテンツの再生側の装置、すなわち音声処理装置に伝送されているものとする。

[0240] フラグ情報hasGainCompensObjectsの値が、聴覚特性テーブルを用いたゲイン補正を行うオブジェクトがある旨の値である場合、ゲイン聴覚特性情報にはオブジェクト数num_objectsの分だけ文字「isGainCompensObject[o]」により示されるフラグ情報が含まれている。

[0241] フラグ情報isGainCompensObject[o]は、o番目のオブジェクトに対して、聴覚特性テーブルを用いたゲイン補正を行うか否かを示している。

[0242] さらに、フラグ情報isGainCompensObject[o]の値が、聴覚特性テーブルを用いたゲイン補正を行う旨の値である場合、ゲイン聴覚特性情報には文字「applyTableIndex[o]」により示されるインデックスが含まれている。

[0243] このインデックスapplyTableIndex[o]は、o番目のオブジェクトに対してゲイン補正を行うときに用いる聴覚特性テーブルを示す情報である。

[0244] 例えば聴覚特性テーブルの数numGainAuditoryPropertyTablesが0である場合、聴覚特性テーブルは伝送されず、ゲイン聴覚特性情報にはインデックスapplyTableIndex[o]も含まれていない。すなわち、インデックスapplyTableIndex[o]は伝送されない。

[0245] そのような場合、例えば聴覚特性テーブル保持部22に保持されている聴覚特性テーブルが用いられてゲイン補正が行われるようにしてもよいし、ゲイン補正が行われないようにしてもよい。

[0246] 〈音声処理装置の構成例〉

以上のようなゲイン聴覚特性情報が符号化ビットストリームにより伝送される場合、音声処理装置は、例えば図18に示すように構成される。なお、

図 1 8 において図 1 4 における場合と対応する部分には同一の符号を付してあり、その説明は適宜省略する。

- [0247] 図 1 8 に示す音声処理装置 1 5 1 は、入力部 1 2 1、位置情報補正部 1 2 2、ゲイン／周波数特性補正部 1 2 3、聴覚特性テーブル保持部 2 2、空間音響特性付加部 1 2 4、レンダラ処理部 1 2 5、および畳み込み処理部 1 2 6 を有している。
- [0248] 音声処理装置 1 5 1 の構成は、図 1 4 に示した音声処理装置 9 1 の構成と同じであるが、符号化ビットストリームから抽出されたゲイン聴覚特性情報から読み出された聴覚特性テーブル等がゲイン／周波数特性補正部 1 2 3 に供給される点で、音声処理装置 9 1 とは異なる。
- [0249] すなわち、音声処理装置 1 5 1 では、ゲイン／周波数特性補正部 1 2 3 にはゲイン聴覚特性情報から読み出された聴覚特性テーブル、フラグ情報 `hasGainCompensObjects`、フラグ情報 `isGainCompensObject[o]`、インデックス `applyTableIndex[o]` などが供給される。
- [0250] 音声処理装置 1 5 1 においては、基本的には図 1 5 を参照して説明した再生信号生成処理が行われる。
- [0251] 但し、ステップ S 7 3 ではゲイン／周波数特性補正部 1 2 3 は、聴覚特性テーブルの数 `numGainAuditoryPropertyTables` が 0 である場合、つまり外部から聴覚特性テーブルが供給されなかった場合、聴覚特性テーブル保持部 2 2 に保持されている聴覚特性テーブルを用いてゲイン補正を行う。
- [0252] これに対して、ゲイン／周波数特性補正部 1 2 3 は、外部から聴覚特性テーブルが供給された場合、その供給された聴覚特性テーブルを用いてゲイン補正を行う。
- [0253] 具体的には、ゲイン／周波数特性補正部 1 2 3 は、外部から供給された複数の聴覚特性テーブルのうち、インデックス `applyTableIndex[o]` により示される聴覚特性テーブルを用いて、`o` 番目のオブジェクトに対するゲイン補正を行う。
- [0254] 但し、ゲイン／周波数特性補正部 1 2 3 は、フラグ情報 `isGainCompensObje`

ct[o]の値が、聴覚特性テーブルを用いたゲイン補正を行わない旨の値であるオブジェクトについては、聴覚特性テーブルを用いたゲイン補正は行わない。

[0255] すなわち、ゲイン／周波数特性補正部 1 2 3 では、聴覚特性テーブルを用いたゲイン補正を行う旨の値のフラグ情報isGainCompensObject[o]が供給された場合、インデックスapplyTableIndex[o]により示される聴覚特性テーブルを用いたゲイン補正が行われる。

[0256] また、例えばゲイン／周波数特性補正部 1 2 3 は、フラグ情報hasGainCompensObjectsの値が、聴覚特性テーブルを用いたゲイン補正を行うオブジェクトがない旨の値である場合には、オブジェクトに対する聴覚特性テーブルを用いたゲイン補正は行わない。

[0257] 以上のように、本技術によれば、オブジェクトオーディオの3Dミキシングや、自由視点のコンテンツの再生などにおいて、各オブジェクトのゲイン情報、すなわちゲイン値を簡単に決定することができる。これにより、より簡単にゲイン補正を行うことができる。

[0258] また、本技術によれば、聴取位置を変更したときの聴取者とオブジェクト（音源）の相対的な位置関係の変化に伴う聴感上の音量の変化を適切に補正することができる。

[0259] 〈コンピュータの構成例〉

ところで、上述した一連の処理は、ハードウェアにより実行することもできるし、ソフトウェアにより実行することもできる。一連の処理をソフトウェアにより実行する場合には、そのソフトウェアを構成するプログラムが、コンピュータにインストールされる。ここで、コンピュータには、専用のハードウェアに組み込まれているコンピュータや、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどが含まれる。

[0260] 図 1 9 は、上述した一連の処理をプログラムにより実行するコンピュータのハードウェアの構成例を示すブロック図である。

- [0261] コンピュータにおいて、CPU (Central Processing Unit) 501, ROM (Read Only Memory) 502, RAM (Random Access Memory) 503は、バス504により相互に接続されている。
- [0262] バス504には、さらに、入出力インターフェース505が接続されている。入出力インターフェース505には、入力部506、出力部507、記録部508、通信部509、及びドライブ510が接続されている。
- [0263] 入力部506は、キーボード、マウス、マイクロホン、撮像素子などよりなる。出力部507は、ディスプレイ、スピーカなどよりなる。記録部508は、ハードディスクや不揮発性のメモリなどよりなる。通信部509は、ネットワークインターフェースなどよりなる。ドライブ510は、磁気ディスク、光ディスク、光磁気ディスク、又は半導体メモリなどのリムーバブル記録媒体511を駆動する。
- [0264] 以上のように構成されるコンピュータでは、CPU501が、例えば、記録部508に記録されているプログラムを、入出力インターフェース505及びバス504を介して、RAM503にロードして実行することにより、上述した一連の処理が行われる。
- [0265] コンピュータ (CPU501) が実行するプログラムは、例えば、パッケージメディア等としてのリムーバブル記録媒体511に記録して提供することができる。また、プログラムは、ローカルエリアネットワーク、インターネット、デジタル衛星放送といった、有線または無線の伝送媒体を介して提供することができる。
- [0266] コンピュータでは、プログラムは、リムーバブル記録媒体511をドライブ510に装着することにより、入出力インターフェース505を介して、記録部508にインストールすることができる。また、プログラムは、有線または無線の伝送媒体を介して、通信部509で受信し、記録部508にインストールすることができる。その他、プログラムは、ROM502や記録部508に、あらかじめインストールしておくことができる。
- [0267] なお、コンピュータが実行するプログラムは、本明細書で説明する順序に

沿って時系列に処理が行われるプログラムであっても良いし、並列に、あるいは呼び出しが行われたとき等の必要なタイミングで処理が行われるプログラムであっても良い。

[0268] また、本技術の実施の形態は、上述した実施の形態に限定されるものではなく、本技術の要旨を逸脱しない範囲において種々の変更が可能である。

[0269] 例えば、本技術は、1つの機能をネットワークを介して複数の装置で分担、共同して処理するクラウドコンピューティングの構成をとることができる。

[0270] また、上述のフローチャートで説明した各ステップは、1つの装置で実行する他、複数の装置で分担して実行することができる。

[0271] さらに、1つのステップに複数の処理が含まれる場合には、その1つのステップに含まれる複数の処理は、1つの装置で実行する他、複数の装置で分担して実行することができる。

[0272] さらに、本技術は、以下の構成とすることも可能である。

[0273] (1)

聴取者から見たオーディオオブジェクトの方向に応じて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正值を決定するゲイン補正值決定部を備える
情報処理装置。

(2)

前記ゲイン補正值決定部は、音の到来方向に対する前記聴取者の3次元の聴覚特性に基づいて前記補正值を決定する

(1)に記載の情報処理装置。

(3)

前記ゲイン補正值決定部は、前記聴取者の向きに基づいて前記補正值を決定する

(1)または(2)に記載の情報処理装置。

(4)

前記ゲイン補正值決定部は、前記オーディオオブジェクトが前記聴取者の後方にある場合、前記オーディオオブジェクトが前記聴取者の前方にある場合よりも前記補正值が大きくなるように前記補正值を決定する

(1) 乃至 (3) の何れか一項に記載の情報処理装置。

(5)

前記ゲイン補正值決定部は、前記オーディオオブジェクトが前記聴取者の側方にある場合、前記オーディオオブジェクトが前記聴取者の前方にある場合よりも前記補正值が小さくなるように前記補正值を決定する

(1) 乃至 (4) の何れか一項に記載の情報処理装置。

(6)

前記ゲイン補正值決定部は、所定の前記方向に応じた前記補正值を、他の方向に応じた前記補正值に基づく補間処理により求めることで、前記所定の前記方向に応じた前記補正值を決定する

(1) 乃至 (5) の何れか一項に記載の情報処理装置。

(7)

前記ゲイン補正值決定部は、前記補間処理としてVBAPを行う

(6) に記載の情報処理装置。

(8)

前記ゲイン補正值決定部は、リニア値またはデシベル値で前記補正值を求める

(7) に記載の情報処理装置。

(9)

情報処理装置が、

聴取者から見たオーディオオブジェクトの方向に応じて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正值を決定する

情報処理方法。

(10)

聴取者から見たオーディオオブジェクトの方向に応じて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正値を決定する

ステップを含む処理をコンピュータに実行させるプログラム。

(11)

オーディオオブジェクトの位置を示す位置情報に基づいて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正値であって、聴取者から見た前記オーディオオブジェクトの方向に応じた補正値を決定し、前記補正値により補正された前記ゲイン値に基づいて前記オーディオ信号の前記ゲイン補正を行うゲイン補正部と、

前記ゲイン補正により得られた前記オーディオ信号に基づいてレンダリング処理を行い、前記オーディオオブジェクトの音を再生するための複数のチャンネルの再生信号を生成するレンダラ処理部と

を備える再生装置。

(12)

前記ゲイン補正部は、前記オーディオ信号のメタデータに含まれている前記ゲイン値を前記補正値により補正する

(11)に記載の再生装置。

(13)

前記ゲイン補正部は、前記ゲイン値の補正を行う旨のフラグが供給された場合、前記補正値により前記ゲイン値を補正する

(11)または(12)に記載の再生装置。

(14)

前記ゲイン補正部は、前記聴取者から見た前記オーディオオブジェクトの方向と、前記補正値とが対応付けられた複数のテーブルのうち、供給されたインデックスにより示される前記テーブルを用いて前記補正値を決定する

(13)に記載の再生装置。

(15)

前記聴取者の位置を示す情報に基づいて、前記オーディオ信号のメタデータに含まれている前記位置情報を補正する位置情報補正部をさらに備え、

前記ゲイン補正部は、補正された前記位置情報に基づいて前記補正値を決定する

(11)乃至(14)の何れか一項に記載の再生装置。

(16)

前記位置情報補正部は、前記聴取者の位置を示す情報、および前記聴取者の向きを示す方向情報に基づいて前記位置情報を補正する

(15)に記載の再生装置。

(17)

再生装置が、

オーディオオブジェクトの位置を示す位置情報に基づいて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正値であって、聴取者から見た前記オーディオオブジェクトの方向に応じた補正値を決定し、

前記補正値により補正された前記ゲイン値に基づいて前記オーディオ信号の前記ゲイン補正を行い、

前記ゲイン補正により得られた前記オーディオ信号に基づいてレンダリング処理を行い、前記オーディオオブジェクトの音を再生するための複数のチャンネルの再生信号を生成する

再生方法。

(18)

オーディオオブジェクトの位置を示す位置情報に基づいて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正値であって、聴取者から見た前記オーディオオブジェクトの方向に応じた補正値を決定し、

前記補正値により補正された前記ゲイン値に基づいて前記オーディオ信号の前記ゲイン補正を行い、

前記ゲイン補正により得られた前記オーディオ信号に基づいてレンダリング処理を行い、前記オーディオオブジェクトの音を再生するための複数のチャンネルの再生信号を生成する

ステップを含む処理をコンピュータに実行させるプログラム。

符号の説明

[0274] 1 1 情報処理装置, 2 1 ゲイン補正值決定部, 2 2 聴覚特性テーブル保持部, 6 2 聴覚特性テーブル生成部, 6 4 表示制御部, 1 2 2 位置情報補正部, 1 2 3 ゲイン／周波数特性補正部

請求の範囲

- [請求項1] 聴取者から見たオーディオオブジェクトの方向に応じて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正値を決定するゲイン補正値決定部を備える情報処理装置。
- [請求項2] 前記ゲイン補正値決定部は、音の到来方向に対する前記聴取者の3次元の聴覚特性に基づいて前記補正値を決定する請求項1に記載の情報処理装置。
- [請求項3] 前記ゲイン補正値決定部は、前記聴取者の向きに基づいて前記補正値を決定する請求項1に記載の情報処理装置。
- [請求項4] 前記ゲイン補正値決定部は、前記オーディオオブジェクトが前記聴取者の後方にある場合、前記オーディオオブジェクトが前記聴取者の前方にある場合よりも前記補正値が大きくなるように前記補正値を決定する請求項1に記載の情報処理装置。
- [請求項5] 前記ゲイン補正値決定部は、前記オーディオオブジェクトが前記聴取者の側方にある場合、前記オーディオオブジェクトが前記聴取者の前方にある場合よりも前記補正値が小さくなるように前記補正値を決定する請求項1に記載の情報処理装置。
- [請求項6] 前記ゲイン補正値決定部は、所定の前記方向に応じた前記補正値を、他の方向に応じた前記補正値に基づく補間処理により求めることで、前記所定の前記方向に応じた前記補正値を決定する請求項1に記載の情報処理装置。
- [請求項7] 前記ゲイン補正値決定部は、前記補間処理としてVBAPを行う請求項6に記載の情報処理装置。
- [請求項8] 前記ゲイン補正値決定部は、リニア値またはデシベル値で前記補正

値を求める

請求項 7 に記載の情報処理装置。

[請求項9]

情報処理装置が、

聴取者から見たオーディオオブジェクトの方向に応じて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正値を決定する

情報処理方法。

[請求項10]

聴取者から見たオーディオオブジェクトの方向に応じて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正値を決定する

ステップを含む処理をコンピュータに実行させるプログラム。

[請求項11]

オーディオオブジェクトの位置を示す位置情報に基づいて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正値であって、聴取者から見た前記オーディオオブジェクトの方向に応じた補正値を決定し、前記補正値により補正された前記ゲイン値に基づいて前記オーディオ信号の前記ゲイン補正を行うゲイン補正部と、

前記ゲイン補正により得られた前記オーディオ信号に基づいてレンダリング処理を行い、前記オーディオオブジェクトの音を再生するための複数のチャンネルの再生信号を生成するレンダラ処理部と

を備える再生装置。

[請求項12]

前記ゲイン補正部は、前記オーディオ信号のメタデータに含まれている前記ゲイン値を前記補正値により補正する

請求項 1 1 に記載の再生装置。

[請求項13]

前記ゲイン補正部は、前記ゲイン値の補正を行う旨のフラグが供給された場合、前記補正値により前記ゲイン値を補正する

請求項 1 1 に記載の再生装置。

[請求項14]

前記ゲイン補正部は、前記聴取者から見た前記オーディオオブジェ

クトの方向と、前記補正值とが対応付けられた複数のテーブルのうち、供給されたインデックスにより示される前記テーブルを用いて前記補正值を決定する

請求項 1 3 に記載の再生装置。

[請求項15] 前記聴取者の位置を示す情報に基づいて、前記オーディオ信号のメタデータに含まれている前記位置情報を補正する位置情報補正部をさらに備え、

前記ゲイン補正部は、補正された前記位置情報に基づいて前記補正值を決定する

請求項 1 1 に記載の再生装置。

[請求項16] 前記位置情報補正部は、前記聴取者の位置を示す情報、および前記聴取者の向きを示す方向情報に基づいて前記位置情報を補正する

請求項 1 5 に記載の再生装置。

[請求項17] 再生装置が、

オーディオオブジェクトの位置を示す位置情報に基づいて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正值であって、聴取者から見た前記オーディオオブジェクトの方向に応じた補正值を決定し、

前記補正值により補正された前記ゲイン値に基づいて前記オーディオ信号の前記ゲイン補正を行い、

前記ゲイン補正により得られた前記オーディオ信号に基づいてレンダリング処理を行い、前記オーディオオブジェクトの音を再生するための複数のチャンネルの再生信号を生成する

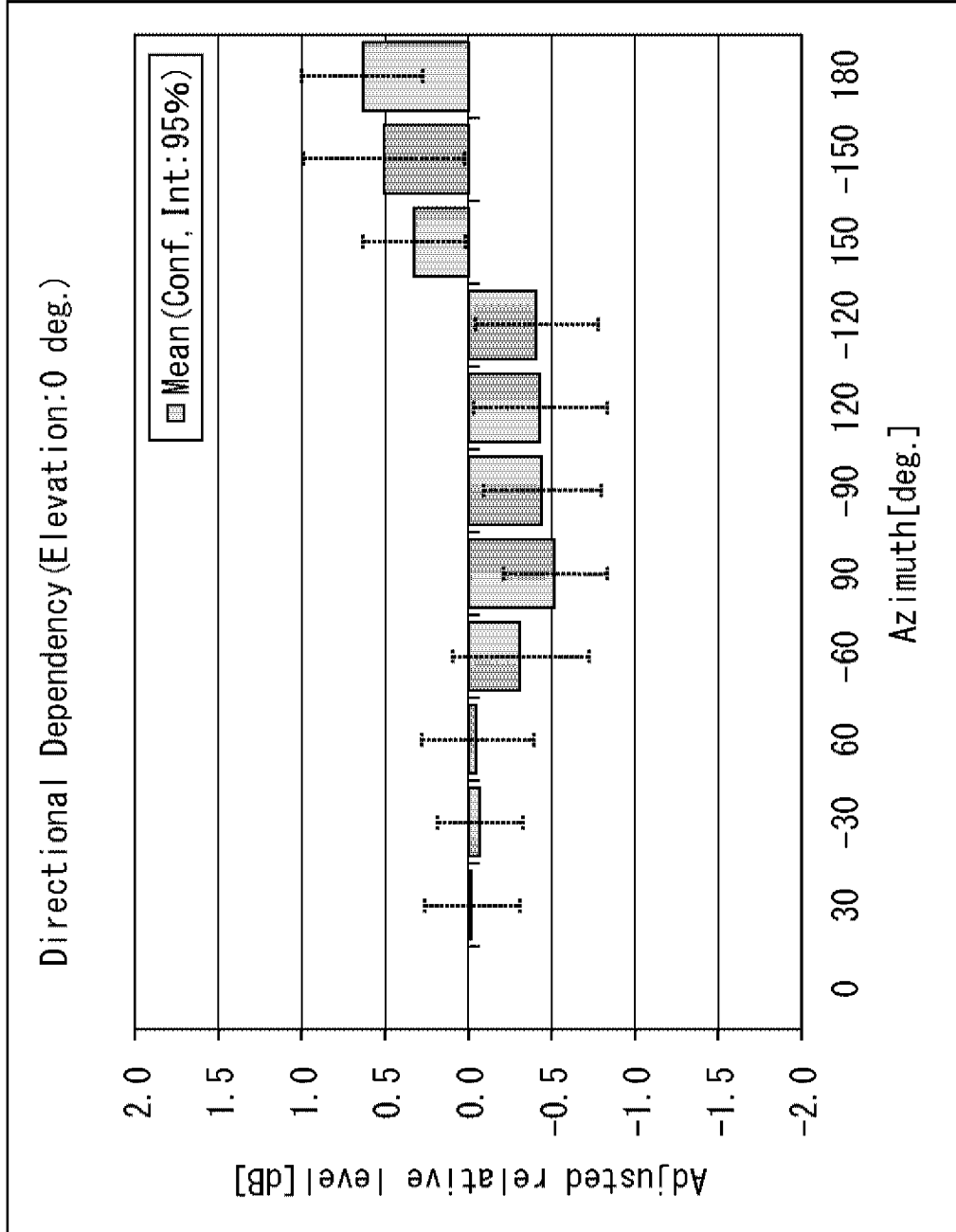
再生方法。

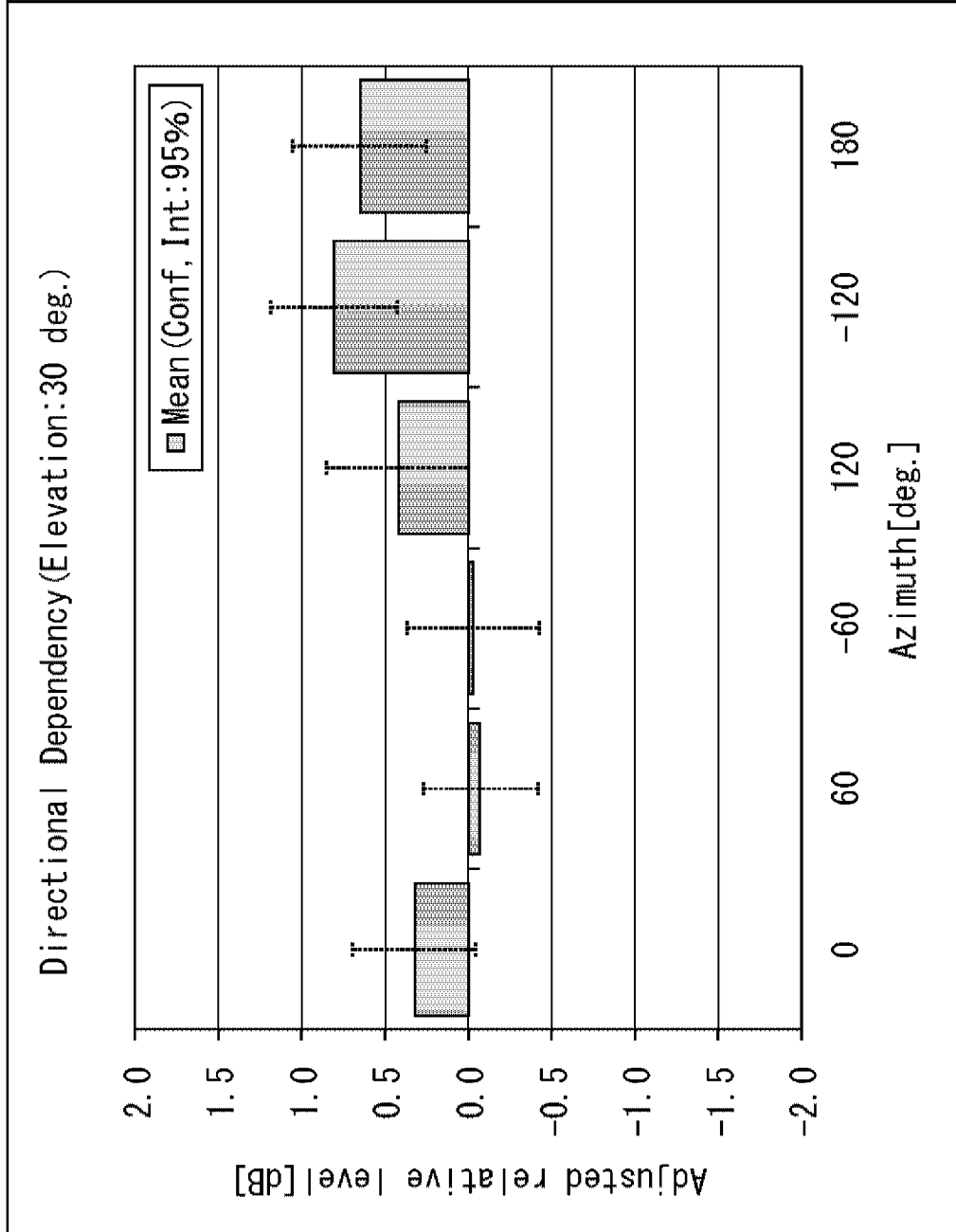
[請求項18] オーディオオブジェクトの位置を示す位置情報に基づいて、前記オーディオオブジェクトのオーディオ信号をゲイン補正するためのゲイン値の補正值であって、聴取者から見た前記オーディオオブジェクトの方向に応じた補正值を決定し、

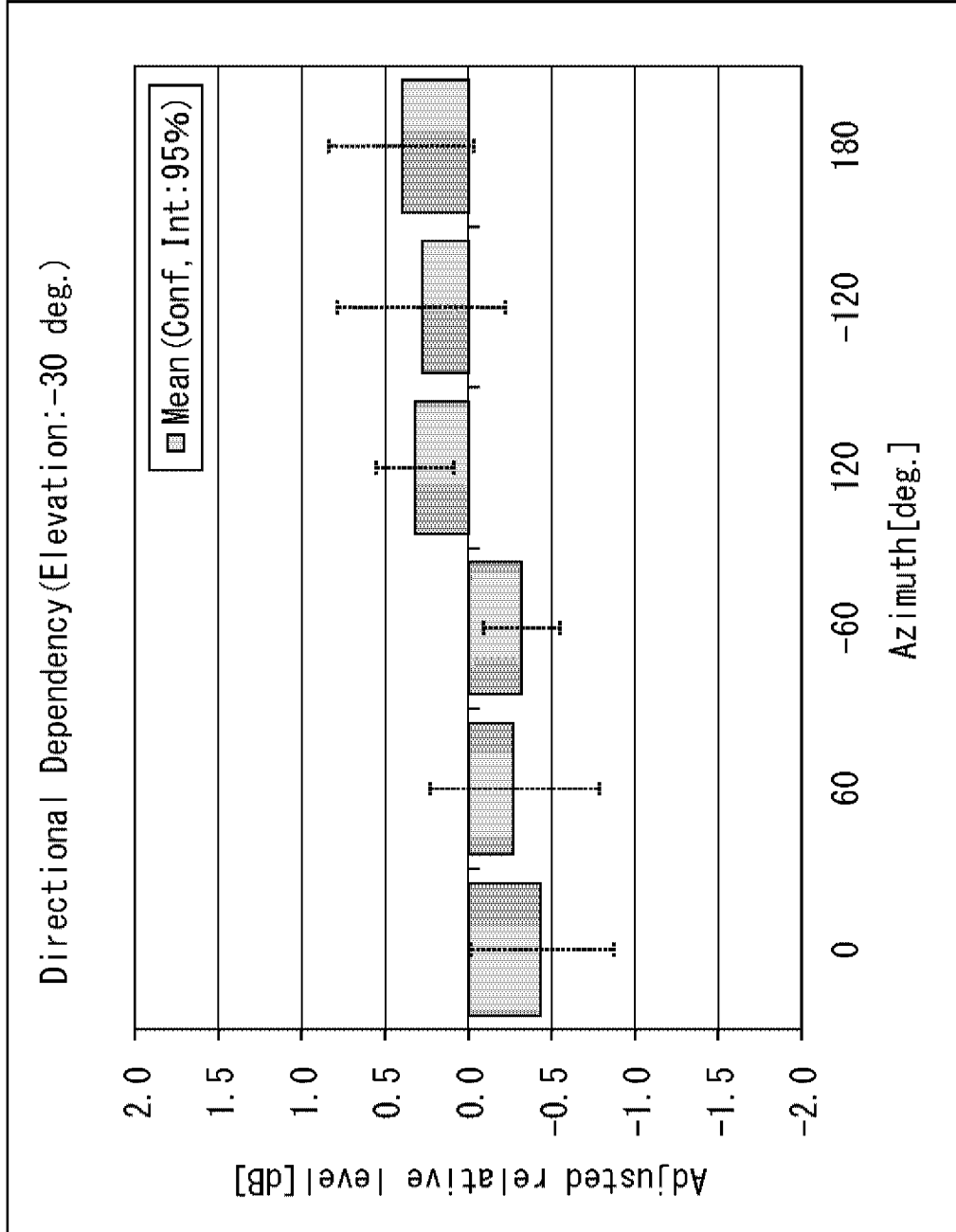
前記補正值により補正された前記ゲイン値に基づいて前記オーディオ信号の前記ゲイン補正を行い、

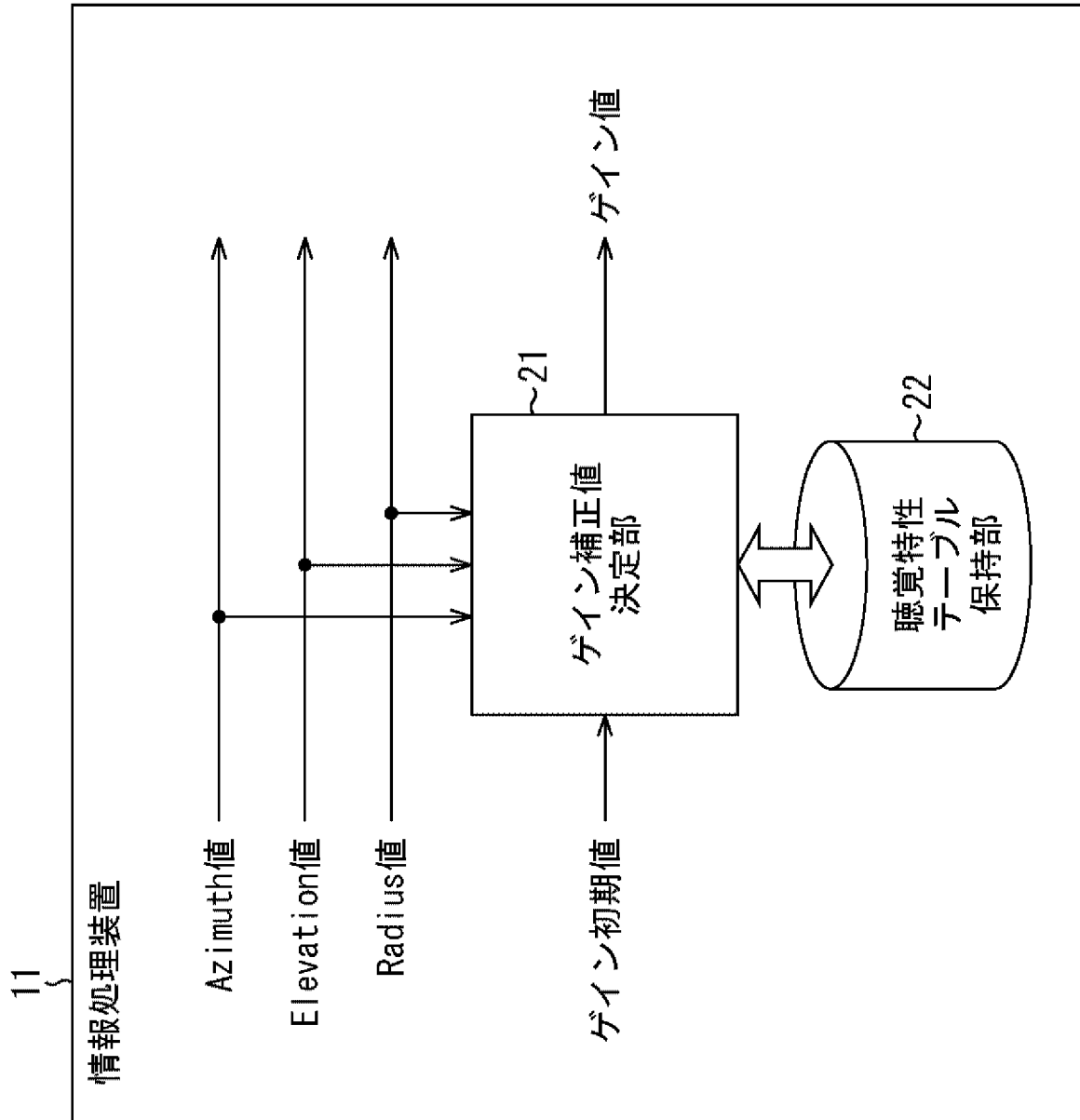
前記ゲイン補正により得られた前記オーディオ信号に基づいてレンダリング処理を行い、前記オーディオオブジェクトの音を再生するための複数のチャンネルの再生信号を生成する

ステップを含む処理をコンピュータに実行させるプログラム。

[図1]
FIG. 1

[図2]
FIG. 2

[図3]
FIG. 3

[図4]
FIG. 4

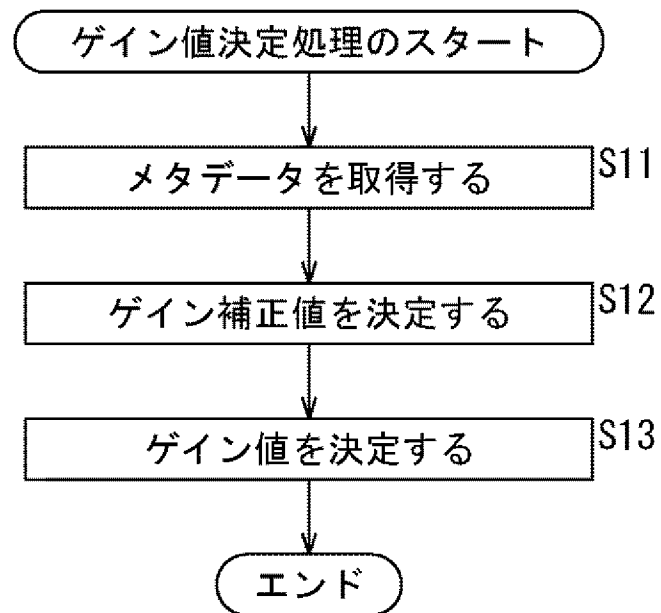
[図5]
FIG. 5

Azimuth [度]	Elevation [度]	Radius [m]	ゲイン補正值 [dB]
0	0	1.0	0.0
30	0	1.0	-0.02
-30	0	1.0	-0.07
60	0	1.0	-0.06
-60	0	1.0	-0.32
90	0	1.0	-0.52
-90	0	1.0	-0.44
120	0	1.0	-0.43
-120	0	1.0	-0.41
150	0	1.0	0.33
-150	0	1.0	0.51
180	0	1.0	0.63

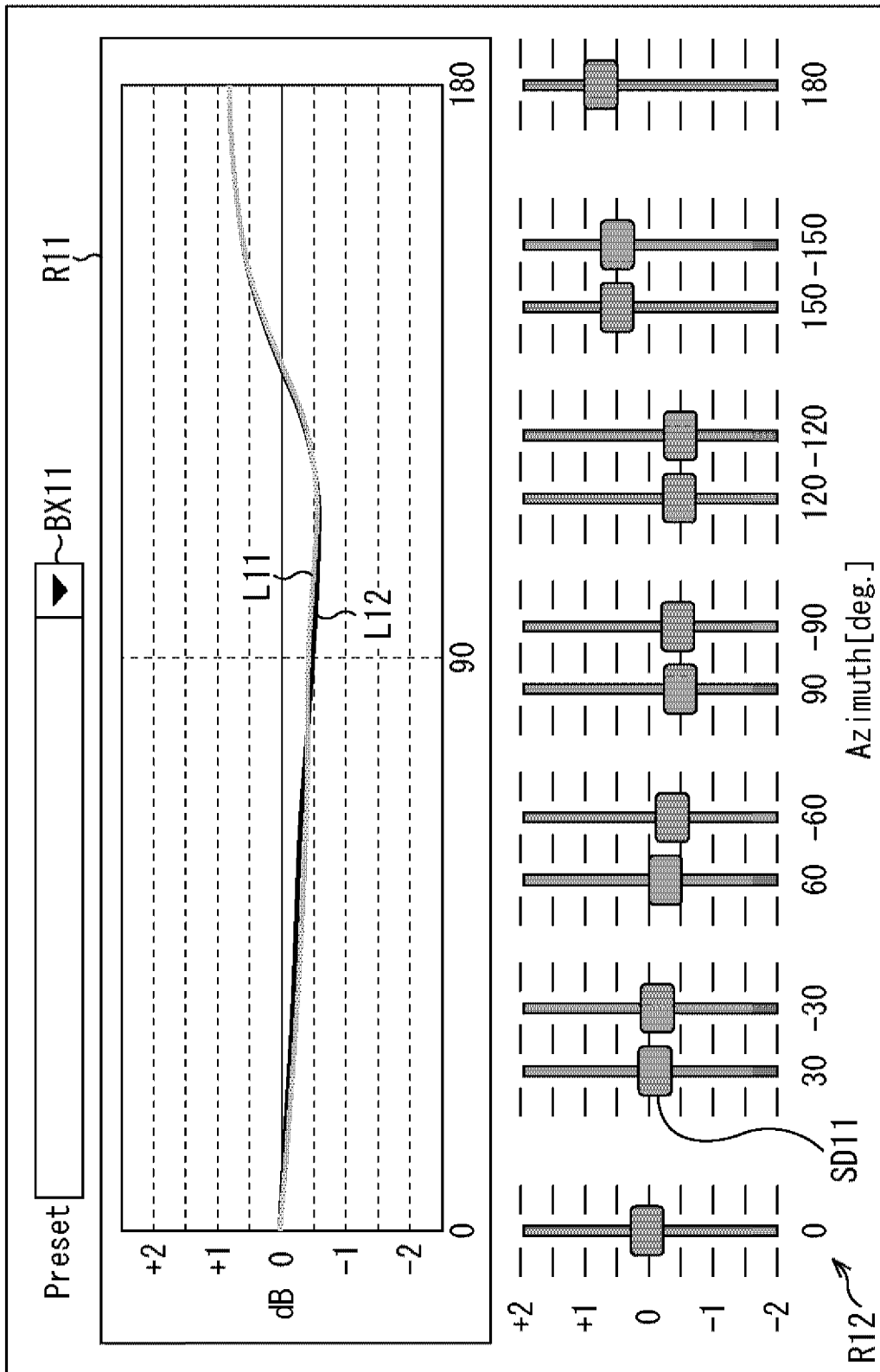
[図6]
FIG. 6

ゲイン補正值 [dB]		Azimuth [度]											
		0	30	-30	60	-60	90	-90	120	-120	150	-150	180
Elevation [度]	30	0.33	-	-	-0.07	-0.03		0.43	0.81	-	-	-	0.66
	0	0.00	-0.02	-0.07	-0.06	-0.32	-0.44	-0.43	-0.41	0.33	0.51	0.63	
	-30	-0.44			-0.28	-0.32		0.32	0.28	-	-	0.40	

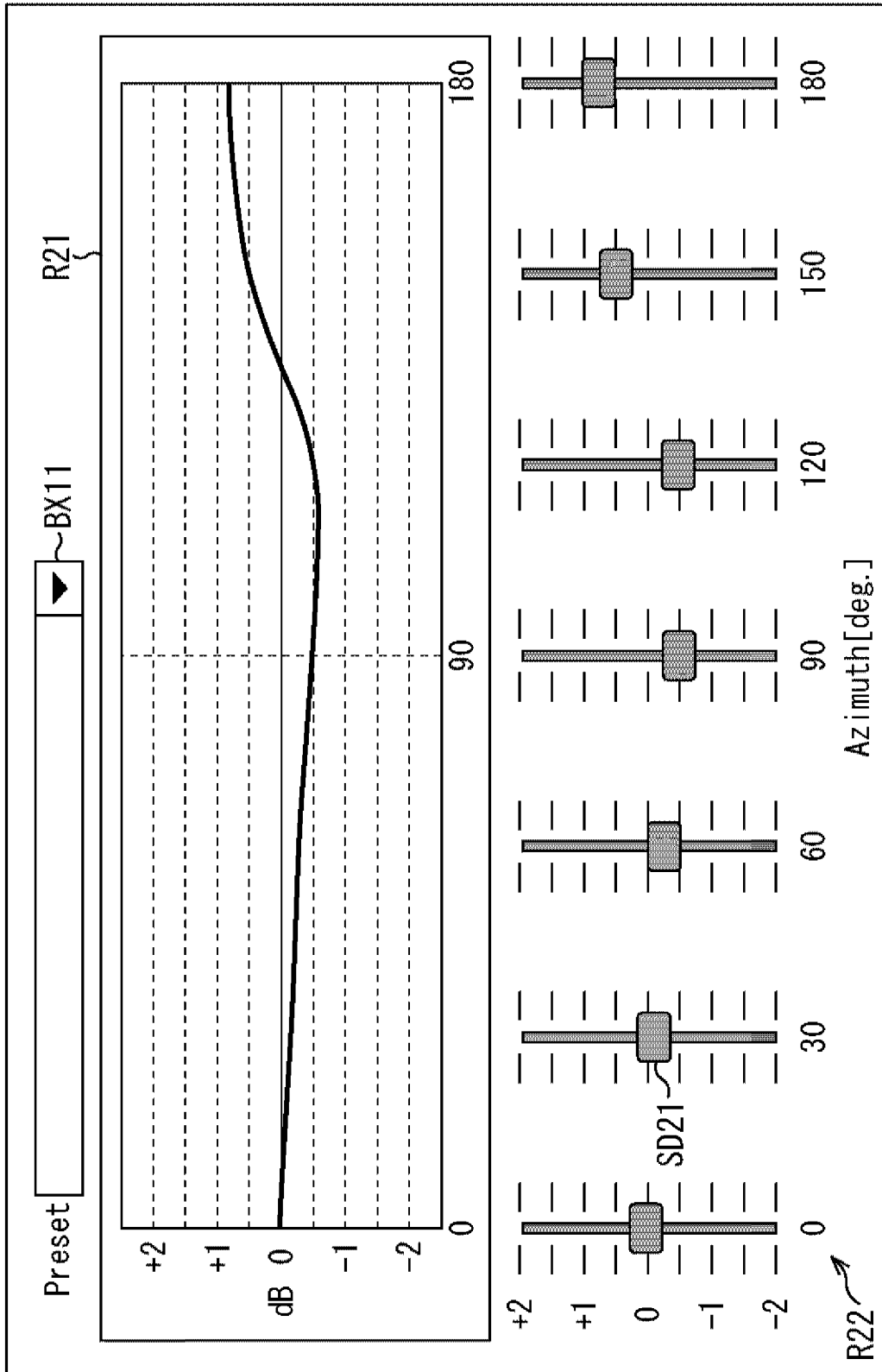
[図7]
FIG. 7



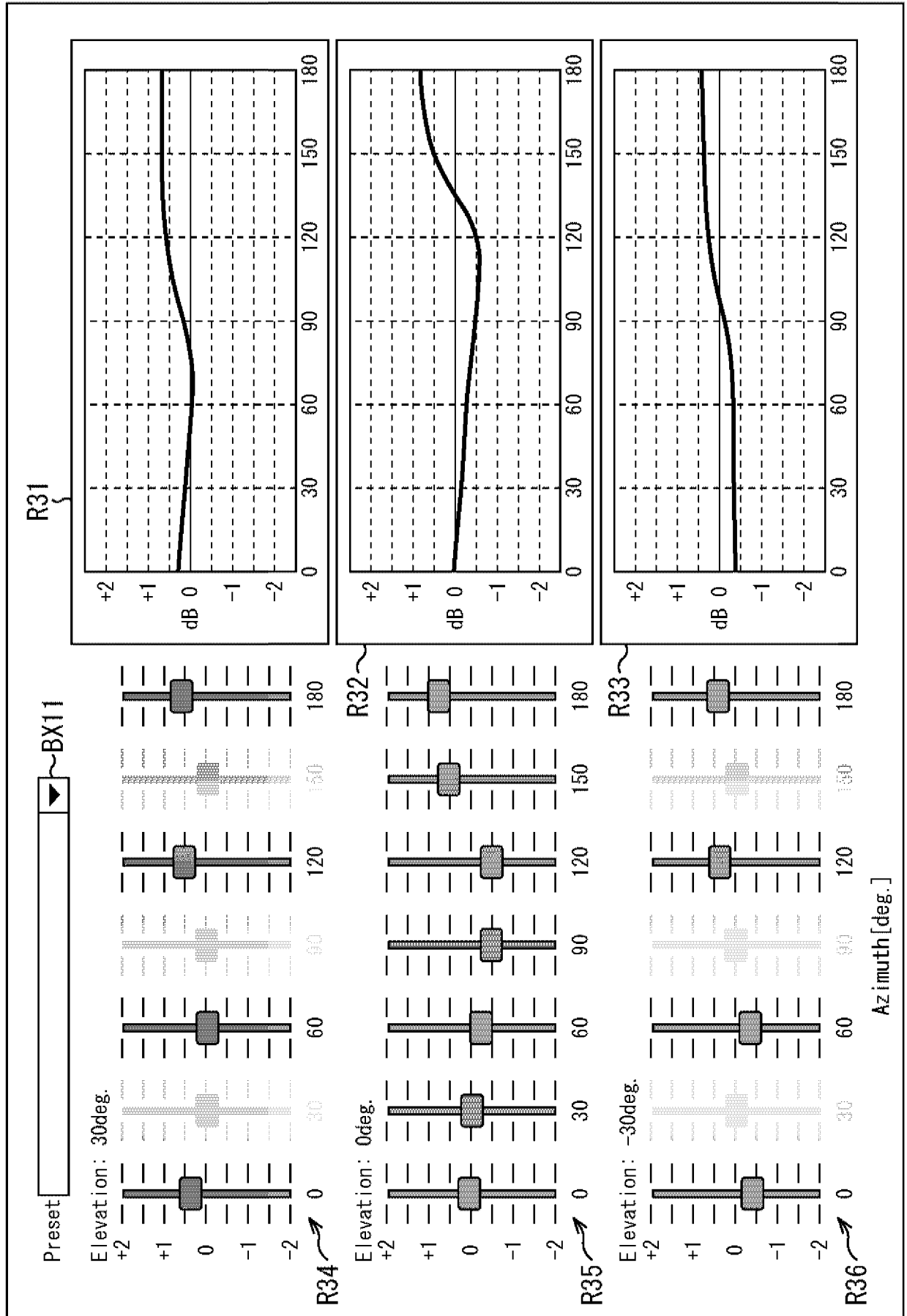
[図8]
FIG. 8



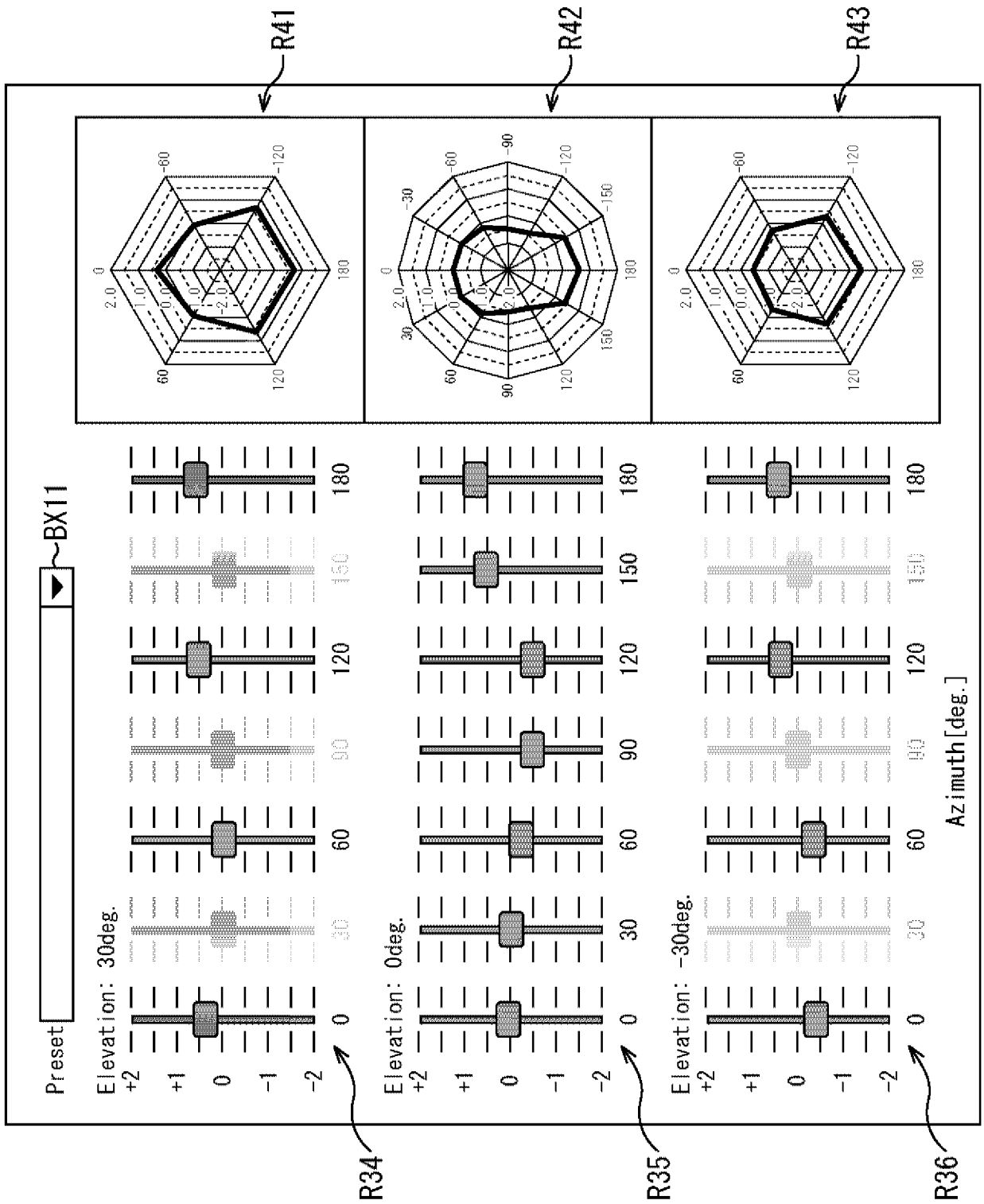
[図9]
FIG. 9



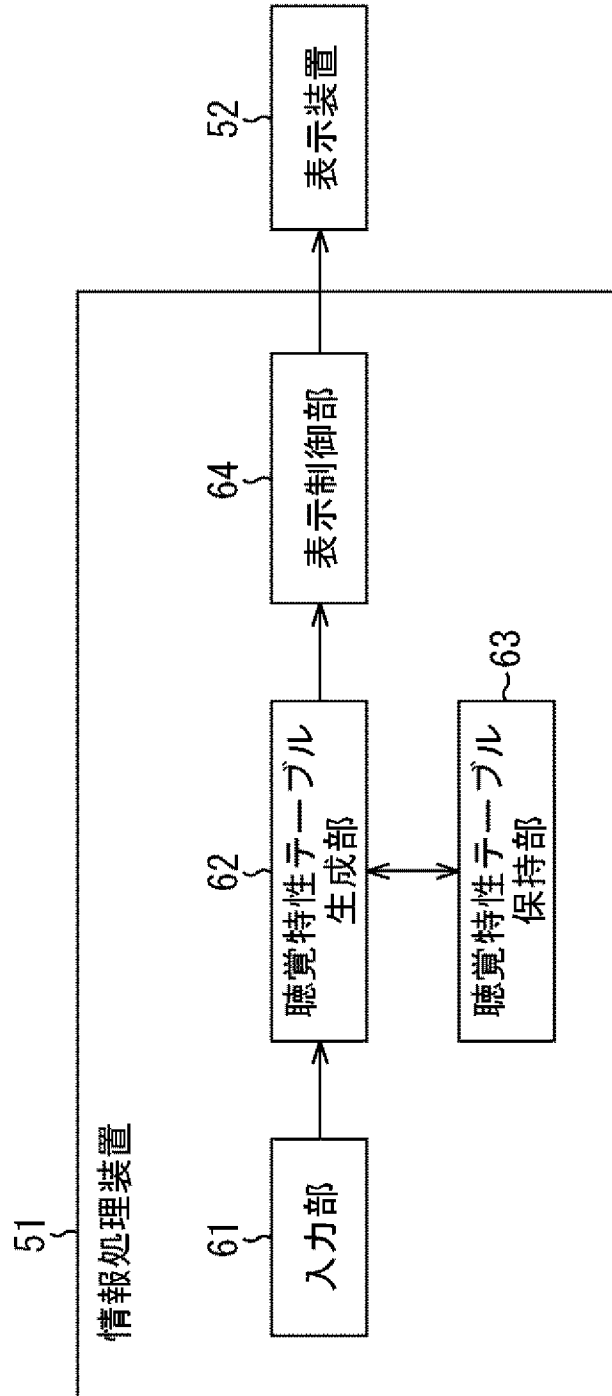
[図10]
FIG. 10



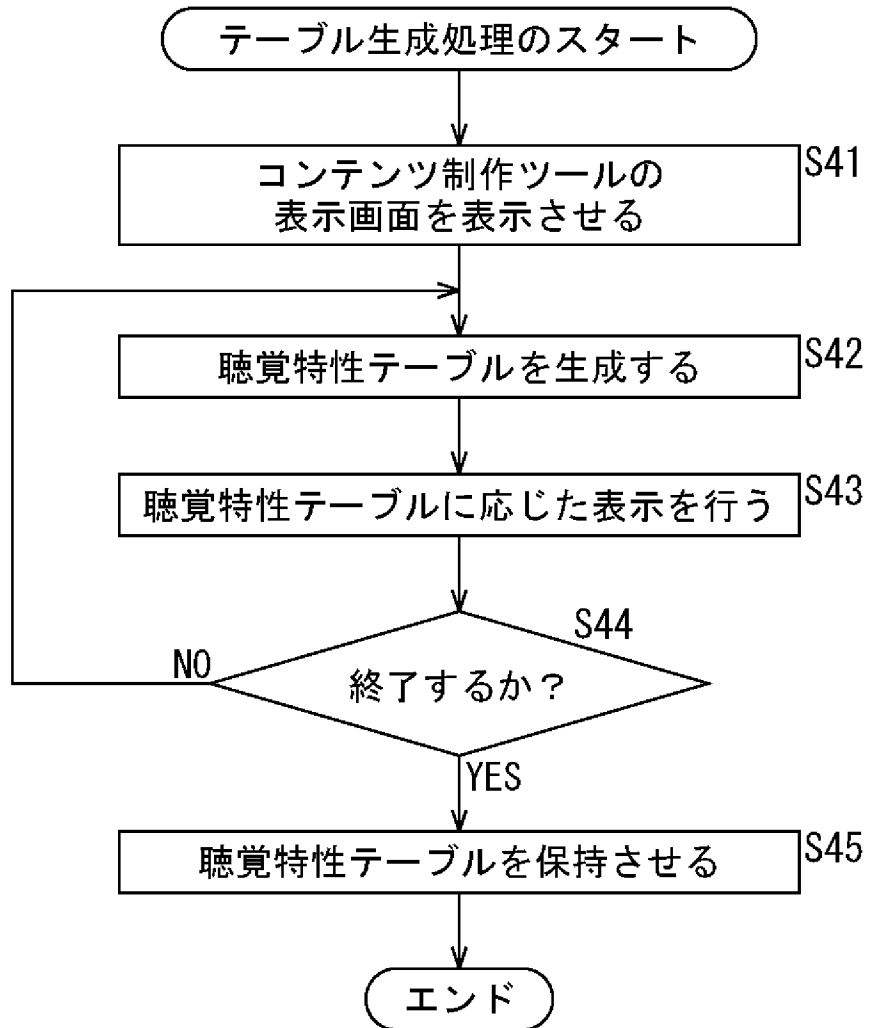
[11]
FIG. 11



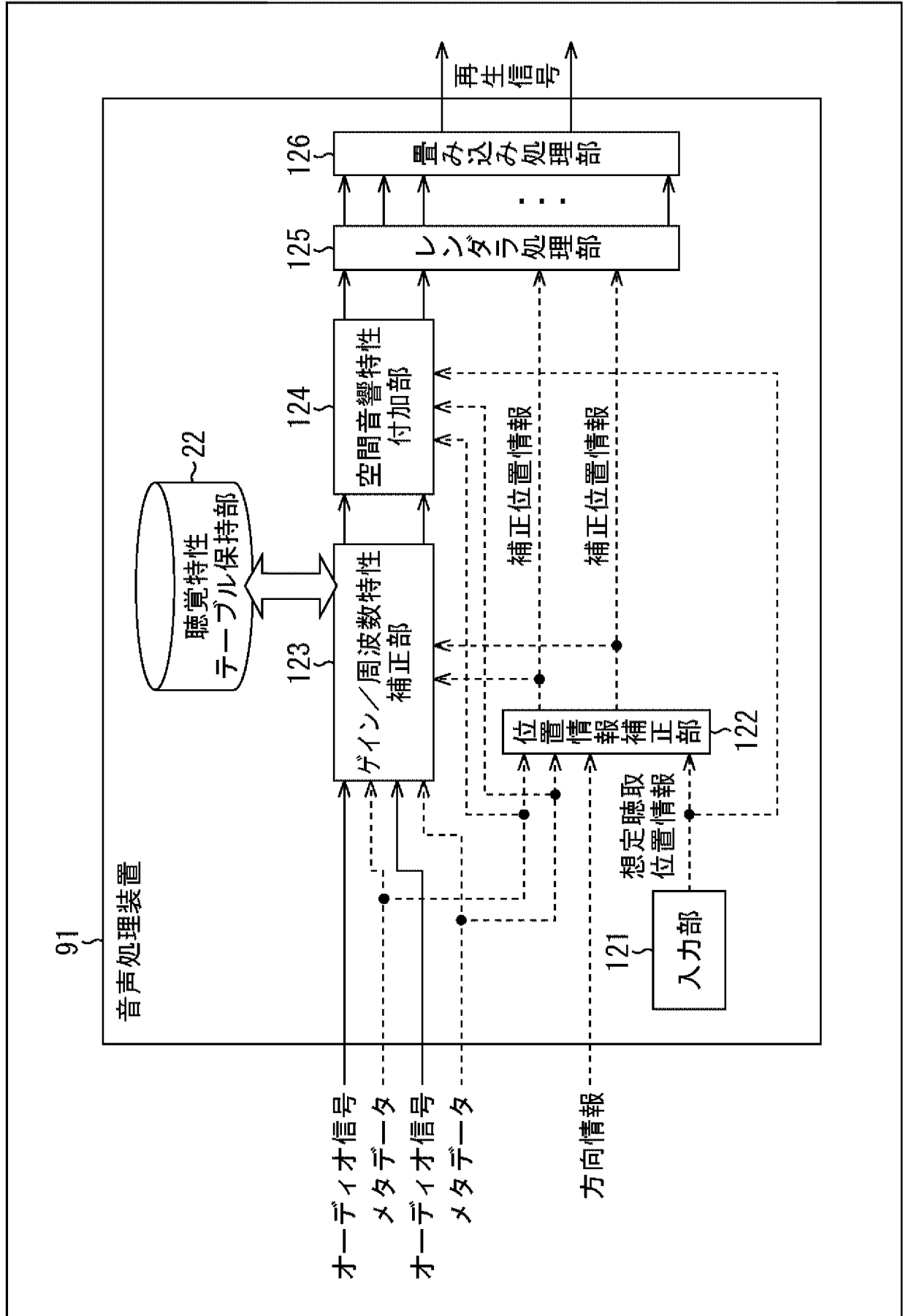
[図12]
FIG. 12



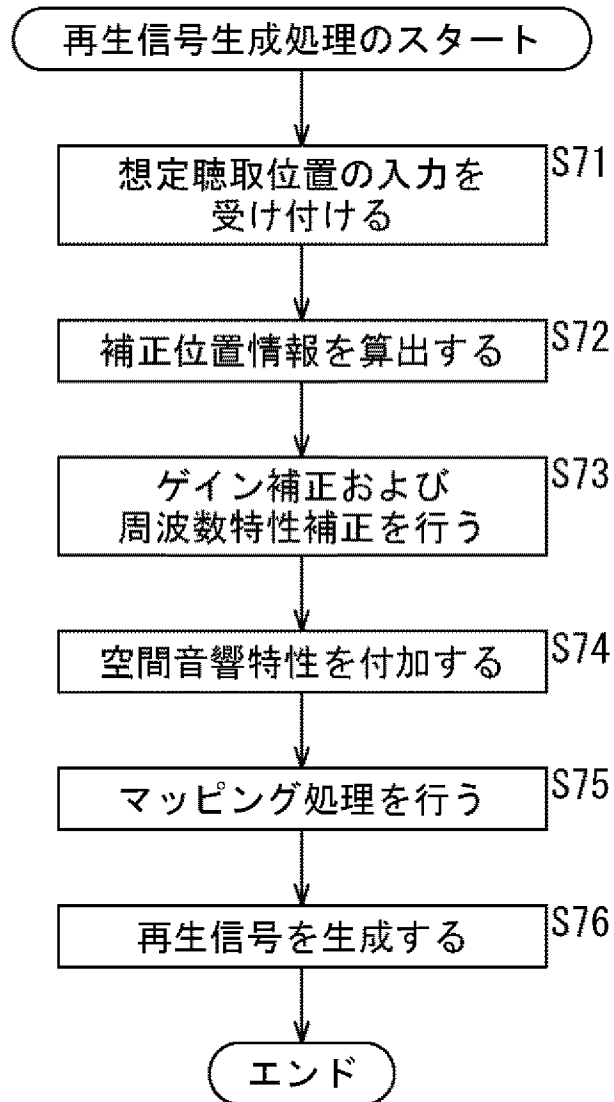
[図13]
FIG. 13



[図14]
FIG. 14



[図15]
FIG. 15



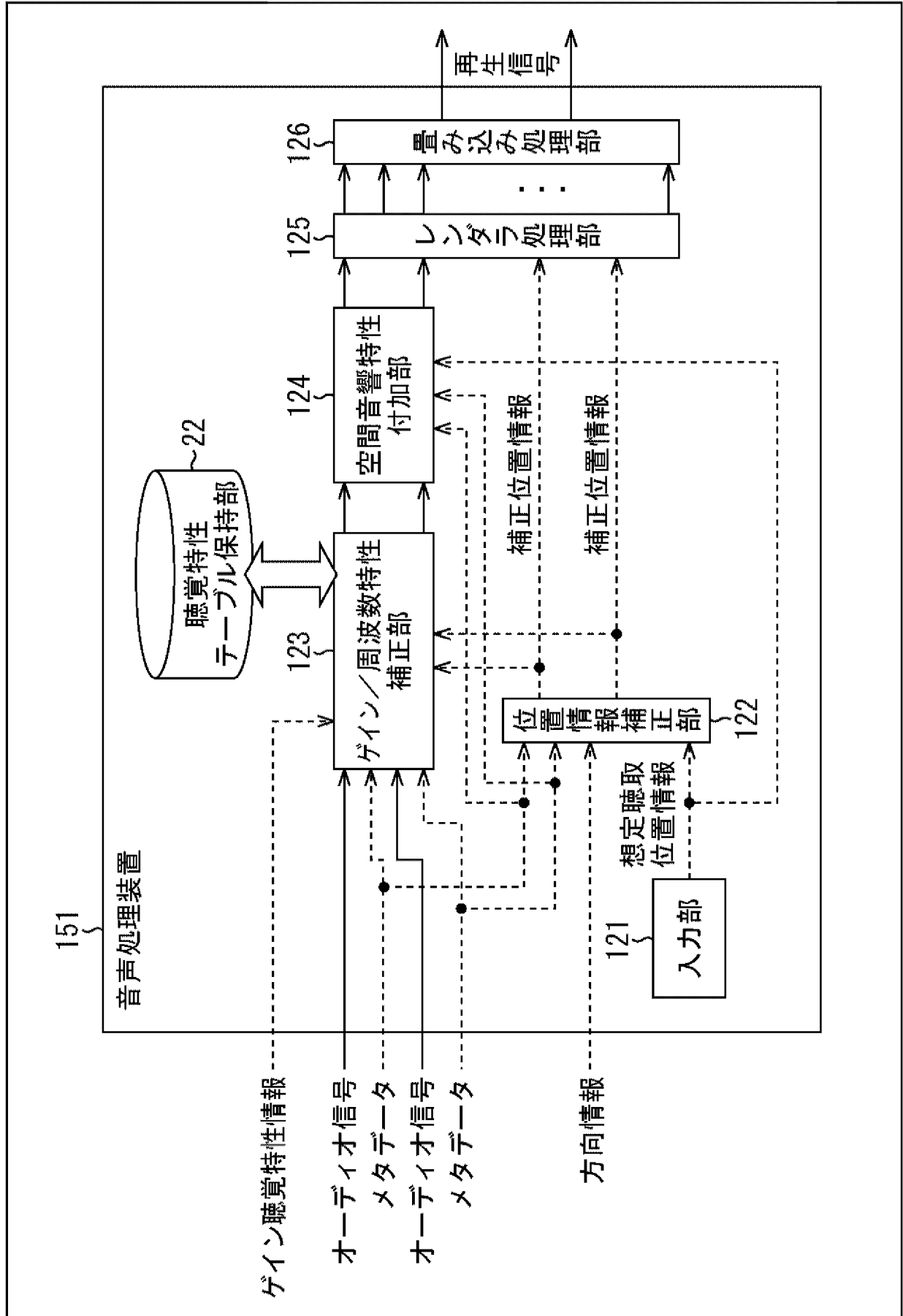
[図16]
FIG. 16

ゲイン補正値 [dB]		Azimuth [度]											
		0	30	-30	60	-60	90	-90	120	-120	150	-150	180
Elevation [度]	30	-0.33	-	-	0.07	0.03			-0.43	-0.81	-	-	-0.66
	0	0.00	0.02	0.07	0.06	0.32	0.44	0.43	0.41	-0.33	-0.51	-0.63	
	-30	0.44			0.28	0.32		-0.32	-0.28	-	-	-0.40	

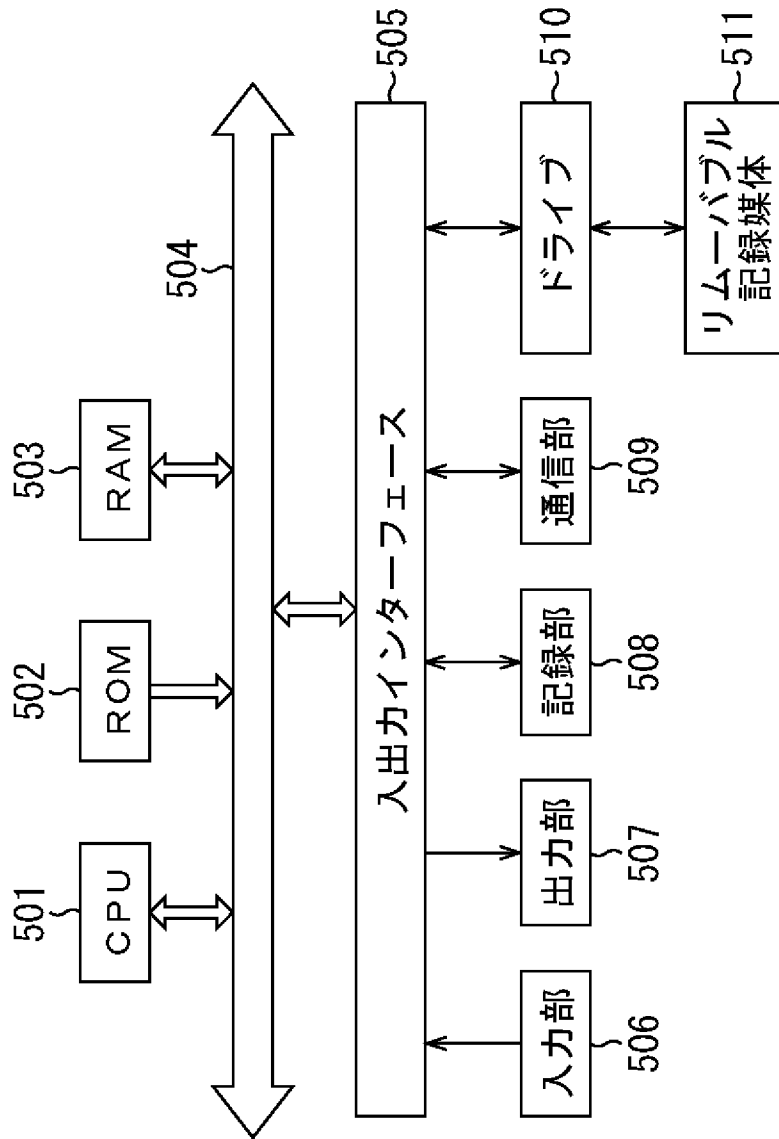
[17]
FIG. 17

Syntax	No. of bits	Mnemonic
<pre> GainAuditoryProperty() { numGainAuditoryPropertyTables for (i = 0; i < numGainAuditoryPropertyTables; i++) { numElements[i] for (n = 0; n < numElements; n++) { azimuth[i][n] elevation[i][n] radius[i][n] gainCompensValue[i][n] } } hasGainCompensObjects; if (hasGainCompensObjects) { for (o = 0; o < num_objects; o++) { isGainCompensObject[o]; if (isGainCompensObject[o] && numGainAuditoryPropertyTables) { applyTableIndex[o]; } } } } </pre>	<p>4</p> <p>4</p> <p>8</p> <p>6</p> <p>4</p> <p>7</p> <p>1</p> <p>1</p> <p>4</p>	<p>uimsbf</p> <p>uimsbf</p> <p>tcimsbf</p> <p>tcimsbf</p> <p>uimsbf</p> <p>tcimsbf</p> <p>bslbf</p> <p>bslbf</p> <p>uimsbf</p>

[図18]
FIG. 18



[図19]
FIG. 19



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2020/014120

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl. H04R3/00 (2006.01) i, H04S5/00 (2006.01) i, H04S7/00 (2006.01) i
 FI: H04S5/00, H04S7/00 310, H04R3/00 310

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int. Cl. H04R3/00, H04S5/00, H04S7/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996
 Published unexamined utility model applications of Japan 1971-2020
 Registered utility model specifications of Japan 1996-2020
 Published registered utility model applications of Japan 1994-2020

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2015/107926 A1 (SONY CORP.) 23 July 2015, paragraphs [0024]-[0124], fig. 1-5	1-3, 6-18
A	JP 2015-126359 A (YAMAHA CORP.) 06 July 2015, paragraphs [0079]-[0092], fig. 11, 12	2, 13
A	WO 2018/096954 A1 (SONY CORP.) 31 May 2018, entire text, all drawings	1-18
A	JP 2018-116299 A (SONY CORP.) 26 July 2018, entire text, all drawings	1-18

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search
02.06.2020

Date of mailing of the international search report
16.06.2020

Name and mailing address of the ISA/
 Japan Patent Office
 3-4-3, Kasumigaseki, Chiyoda-ku,
 Tokyo 100-8915, Japan

Authorized officer

 Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2020/014120

Patent Documents referred to in the Report	Publication Date	Patent Family	Publication Date
WO 2015/107926 A1	23.07.2015	US 2016/0337777 A1 paragraphs [0031]- [0124], fig. 1-5	
JP 2015-126359 A	06.07.2015	US 2015/0237446 A1 paragraphs [0211]- [0220], fig. 22, 23	
WO 2018/096954 A1	31.05.2018	EP 3547718 A1 entire text, all drawings	
JP 2018-116299 A	26.07.2018	US 2019/0130922 A1 entire text, all drawings	

<p>A. 発明の属する分野の分類（国際特許分類（IPC）） H04R 3/00(2006.01)i; H04S 5/00(2006.01)i; H04S 7/00(2006.01)i FI: H04S5/00; H04S7/00 310; H04R3/00 310</p>										
<p>B. 調査を行った分野</p>										
<p>調査を行った最小限資料（国際特許分類（IPC）） H04R3/00; H04S5/00; H04S7/00</p>										
<p>最小限資料以外の資料で調査を行った分野に含まれるもの</p> <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922 - 1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971 - 2020年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996 - 2020年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994 - 2020年</td> </tr> </table>			日本国実用新案公報	1922 - 1996年	日本国公開実用新案公報	1971 - 2020年	日本国実用新案登録公報	1996 - 2020年	日本国登録実用新案公報	1994 - 2020年
日本国実用新案公報	1922 - 1996年									
日本国公開実用新案公報	1971 - 2020年									
日本国実用新案登録公報	1996 - 2020年									
日本国登録実用新案公報	1994 - 2020年									
<p>国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）</p>										
<p>C. 関連すると認められる文献</p>										
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号								
X	WO 2015/107926 A1 (ソニー株式会社) 23.07.2015 (2015 - 07 - 23) [0024]-[0124], 図1-5	1-3, 6-18								
A	JP 2015-126359 A (ヤマハ株式会社) 06.07.2015 (2015 - 07 - 06) [0079]-[0092], 図11-12	2, 13								
A	WO 2018/096954 A1 (ソニー株式会社) 31.05.2018 (2018 - 05 - 31) 全文, 全図	1-18								
A	JP 2018-116299 A (ソニー株式会社) 26.07.2018 (2018 - 07 - 26) 全文, 全図	1-18								
<p><input type="checkbox"/> C欄の続きにも文献が列挙されている。 <input checked="" type="checkbox"/> パテントファミリーに関する別紙を参照。</p>										
* 引用文献のカテゴリー	<p>“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの</p> <p>“A” 特に関連のある文献ではなく、一般的な技術水準を示すもの</p> <p>“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの</p> <p>“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの</p> <p>“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）</p> <p>“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの</p> <p>“O” 口頭による開示、使用、展示等に言及する文献</p> <p>“&” 同一パテントファミリー文献</p> <p>“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献</p>									
国際調査を完了した日	02.06.2020	国際調査報告の発送日 16.06.2020								
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号	権限のある職員（特許庁審査官） 鈴木 圭一郎 5Z 3141 電話番号 03-3581-1101 内線 3591									

国際調査報告
 パテントファミリーに関する情報

国際出願番号

PCT/JP2020/014120

引用文献			公表日	パテントファミリー文献			公表日
WO	2015/107926	A1	23.07.2015	US	2016/0337777	A1	
				[0031]-[0124], figures 1-5			
JP	2015-126359	A	06.07.2015	US	2015/0237446	A1	
				[0211]-[0220], figures 22-23			
WO	2018/096954	A1	31.05.2018	EP	3547718	A1	
				全文, 全図			
JP	2018-116299	A	26.07.2018	US	2019/0130922	A1	
				全文, 全図			